

XX WPI: 1998-240074/21.
DR P-PSDB: AAW56097.
XX
PT New isolated S-adenosyl-L-homocysteine hydrolase enzyme - is used to
PT develop products which can be used in the treatment of e.g.
PT auto-immune disease, transplantations or cancers
XX
PS Claim 7: Fig 1: 33pp; English.
XX
CC This is the nucleotide sequence of the DD4b9.3 enzyme which has
CC S-adenosyl-L-homocysteine hydrolase (AHCY)-type activity. Its
CC products can be used to identify substances which have a stimulatory
CC or inhibitory effect on the enzyme activity. Such substances can be
CC used to modulate dendritic cell (DC) function and for immunomodulation.
CC They can be used in the treatment of e.g. autoimmune diseases,
CC transplantations or cancers. The products can also be used for
CC detection and diagnosis.
XX
SQ Sequence 2563 BP; 646 A; 604 C; 677 G; 636 T; 0 other;
Query Match 100.0%; Score 2563; DB 19; Length 2563;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCGCGGGGAGGCTCGAGCTCGAGCTGCTGCTTCTGTTCTTGTGGCCACCGTGGCT 60
DB 1 GCGCGGGGAGGCTCGAGCTCGAGCTGCTGCTTCTGTTCTTGTGGCCACCGTGGCT 60
QY 61 GTCCGGCTGCTTGGGCTGCGCAACAGAGCGGTGGCCACAGCACCCTAGAACCGA 120
DB 61 GTCCGGCTGCTTGGGCTGCGCAACAGAGCGGTGGCCACAGCACCCTAGAACCGA 120
QY 121 CGCAGCTCGACGAGGGGGCGGAGAGGGTGGCGGATCGCGTGTGGAGGGCGCGCC 180
DB 121 CGCAGCTCGACGAGGGGGCGGAGAGGGTGGCGGATCGCGTGTGGAGGGCGCGCC 180
QY 181 GGGCAGGG 240
DB 181 GGGCAGGG 240
QY 241 GGGCGGGGGGGGAGATGTGATGCTGACGCGGATGCGCGGCGGGGGGGGGGGGG 300
DB 241 GGGCGGGGGGGGAGATGTGATGCTGACGCGGATGCGCGGCGGGGGGGGGGGGG 300
QY 301 TGAAGCAGGCGCAAGAGATCGAGAGCGCGAGAGTACTCTTCATGGCCACCGTCA 360
DB 301 TGAAGCAGGCGCAAGAGATCGAGAGCGCGAGAGTACTCTTCATGGCCACCGTCA 360
QY 361 AGGCGCCCAAGAGCAATCCAGTTTGTGATGACATGACAGAGTTCACCAATTCCCA 420
DB 361 AGGCGCCCAAGAGCAATCCAGTTTGTGATGACATGACAGAGTTCACCAATTCCCA 420
QY 421 CCAAACTGGCGGAAGATCTTGTCTGCTGCTGATCTCAGAGTCTCCACTGACAGCT 480
DB 421 CCAAACTGGCGGAAGATCTTGTCTGCTGCTGATCTCAGAGTCTCCACTGACAGCT 480
QY 481 GTTCAGCTGATCTTACACAGATGCTGATGATGAGTTTCTCCCGAGAGAGAGCAG 540
DB 481 GTTCAGCTGATCTTACACAGATGCTGATGATGAGTTTCTCCCGAGAGAGAGCAG 540
QY 541 AAACCACTCCAAAGGAGGAGCAATTTCTGTGTAAGAACATCAAGCAGAGGAATTG 600
DB 541 AAACCACTCCAAAGGAGGAGCAATTTCTGTGTAAGAACATCAAGCAGAGGAATTG 600
QY 601 GACGCGGGAGATGAGATGACAGAGCAAGACATGTCTGTGATTTCACTCAGGAAC 660
DB 601 GACGCGGGAGATGAGATGACAGAGCAAGACATGTCTGTGATTTCACTCAGGAAC 660
QY 661 GTTCTCAGGGGAGAGCGCTTGGCTGTCTAAATAGTGGCTGTACACATCAGAG 720
DB 661 GTTCTCAGGGGAGAGCGCTTGGCTGTCTAAATAGTGGCTGTACACATCAGAG 720

QY 721 CCCAGACAGCGGTGTGATGAGACACTCTGTGCTGGGGGCTCAGTGGCGCTGCTG 780
DB 721 CCCAGACAGCGGTGTGATGAGACACTCTGTGCTGGGGGCTCAGTGGCGCTGCTG 780
QY 781 CTGTGACATCTACTCAACTCAGATGAGTACCTGCAGCAGCTGGCTGAGGCTGAG 840
DB 781 CTGTGACATCTACTCAACTCAGATGAGTACCTGCAGCAGCTGGCTGAGGCTGAG 840
QY 841 CAGTGTGCTTGAAGGGGAGTCAAGATGACTTCTGTGTGTATGACCGCTGTG 900
DB 841 CAGTGTGCTTGAAGGGGAGTCAAGATGACTTCTGTGTGTATGACCGCTGTG 900
QY 901 TGAACATGATGGGTGGCAGGCCAATGATCTGTGATGATGGGAGACTTAACCACT 960
DB 901 TGAACATGATGGGTGGCAGGCCAATGATCTGTGATGATGGGAGACTTAACCACT 960
QY 961 GGGTTTAAAGAGTATCCAAACGTGTTAAGATCCGAGGCACTGTGGAAGAGCG 1020
DB 961 GGGTTTAAAGAGTATCCAAACGTGTTAAGATCCGAGGCACTGTGGAAGAGCG 1020
QY 1021 TGACTGTGTTTCAAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTCCGCA 1080
DB 1021 TGACTGTGTTTCAAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTCCGCA 1080
QY 1081 TGAACGTCAATGATCTCTGTACCAACAGAGTTGATACTGTACTGCTGCCAGAT 1140
DB 1081 TGAACGTCAATGATCTCTGTACCAACAGAGTTGATACTGTACTGCTGCCAGAT 1140
QY 1141 CCATTTGATGGCTTGAAGAGGACACAGATGATGTTGTTGGGAAACAAGTGTGG 1200
DB 1141 CCATTTGATGGCTTGAAGAGGACACAGATGATGTTGTTGGGAAACAAGTGTGG 1200
QY 1201 TGTGTGCTATGCTGAGGTAGGCAAGGCTGCTGTCTCTCAAGCTCTTGAAGCA 1260
DB 1201 TGTGTGCTATGCTGAGGTAGGCAAGGCTGCTGTCTCTCAAGCTCTTGAAGCA 1260
QY 1261 TTGTCTACATTAACGAAATCGACCCCATCTGTGCTCTGCAAGGCTGCATGATG 1320
DB 1261 TTGTCTACATTAACGAAATCGACCCCATCTGTGCTCTGCAAGGCTGCATGATG 1320
QY 1321 GGGTGTAAAGCTAAATGAAGTATCCGGCAAGTGTCTGTAATACCTGACAGAA 1380
DB 1321 GGGTGTAAAGCTAAATGAAGTATCCGGCAAGTGTCTGTAATACCTGACAGAA 1380
QY 1381 ATAAAGTGTAGTGACACAGGAGCACTTGATGATGAAACAGTTGTATGATGCA 1440
DB 1381 ATAAAGTGTAGTGACACAGGAGCACTTGATGATGAAACAGTTGTATGATGCA 1440
QY 1441 ATATGGCCACTCCAAACAGAAATGATGATGACCAAGCTCCGAGCTGACGT 1500
DB 1441 ATATGGCCACTCCAAACAGAAATGATGATGACCAAGCTCCGAGCTGACGT 1500
QY 1501 GGGAGCAGTACGTTCTCAGGTGACCATGCTGCTGGCCAGATGGCAACGAGTTGCC 1560
DB 1501 GGGAGCAGTACGTTCTCAGGTGACCATGCTGCTGGCCAGATGGCAACGAGTTGCC 1560
QY 1561 TCCTGGCAGAGGGTCTCTACTCAATTTGAGCTCTCCACAGTTCCCACTTGTCTGT 1620
DB 1561 TCCTGGCAGAGGGTCTCTACTCAATTTGAGCTCTCCACAGTTCCCACTTGTCTGT 1620
QY 1621 CCATCAGAGCCACACACAGAGCTTTGGCAGTATGAACTCTATATGACCCGAGGGGC 1680
DB 1621 CCATCAGAGCCACACACAGAGCTTTGGCAGTATGAACTCTATATGACCCGAGGGGC 1680
QY 1681 GATACAGCAGAGTGTACTTGTCTCTTAAGAAATGATGATGATGATGATGATGATG 1740
DB 1681 GATACAGCAGAGTGTACTTGTCTCTTAAGAAATGATGATGATGATGATGATGATG 1740
QY 1741 ATCTGCATCATTTGATGCCCCACCTTACAGAGCTGACAGATGACCAAGCAATACTGG 1800
DB 1741 ATCTGCATCATTTGATGCCCCACCTTACAGAGCTGACAGATGACCAAGCAATACTGG 1800
QY 1801 GACTCAACAAAAATGGGCATTCAAACCTAATTATACAGATACTAATGACCATACTAC 1860

```

|||||
Db 1801 GACTCAACAAAATGGGCATTCAAACCTAATTATACAGATACATAAGACCATACTAC 1860
QY 1861 CAAGACCAGTCCACCTGGAACACACACTCTAAGAATATTTTAAAGATAACTTTAT 1920
Db 1861 CAAGACCAGTCCACCTGGAACACACACTCTAAGAATATTTTAAAGATAACTTTAT 1920
QY 1921 TTCTCTCTACTCCTTCCCTCTGATTTTCTCTATATTCATCTCTGTTTTCATC 1980
Db 1921 TTCTCTCTACTCCTTCCCTCTGATTTTCTCTATATTCATCTCTGTTTTCATC 1980
QY 1981 TCATTATCCAAAGTCTGACAGACACACAGAACTGCTCATGGCTCTTATGATGAATA 2040
Db 1981 TCATTATCCAAAGTCTGACAGACACACAGAACTGCTCATGGCTCTTATGATGAATA 2040
QY 2041 GAAGTCAAGGTCCTCCTACCTCTAGTCACTAAAGAAGATTTACTCCCCAGCCAGAAA 2100
Db 2041 GAAGTCAAGGTCCTCCTACCTCTAGTCACTAAAGAAGATTTACTCCCCAGCCAGAAA 2100
QY 2101 GGTGATCTCTCTTACCATTTCTGGGACTTTAGTCTTAATTAGGTACTTATTAACA 2160
Db 2101 GGTGATCTCTCTTACCATTTCTGGGACTTTAGTCTTAATTAGGTACTTATTAACA 2160
QY 2161 GGAATGCTAAGTACCTCTCTGGAACAATCTGCAATGCTAAATCGCCTTAAAGA 2220
Db 2161 GGAATGCTAAGTACCTCTCTGGAACAATCTGCAATGCTAAATCGCCTTAAAGA 2220
QY 2221 GCCCATTTCTAGCTGCTGAATGCTCTTTCATCTTTCAGAGAGAGGAGTGT 2280
Db 2221 GCCCATTTCTAGCTGCTGAATGCTCTTTCATCTTTCAGAGAGAGGAGTGT 2280
QY 2281 ACCTACCCGGCAGGTAGGTAGATGCGGTGTCATGTTATTTCCCTTAAAGTTCCA 2340
Db 2281 ACCTACCCGGCAGGTAGGTAGATGCGGTGTCATGTTATTTCCCTTAAAGTTCCA 2340
QY 2341 AGCCCTGTTTCTGCTGAAGGTGATGTCAGTTCAGAGATGTATATATGAGCATGG 2400
Db 2341 AGCCCTGTTTCTGCTGAAGGTGATGTCAGTTCAGAGATGTATATATGAGCATGG 2400
QY 2401 CTGTGTAAGATCAGAGAGCCCACTGGAATTTATAGTATAGCCCTTCCCTCCACCTCCA 2460
Db 2401 CTGTGTAAGATCAGAGAGCCCACTGGAATTTATAGTATAGCCCTTCCCTCCACCTCCA 2460
QY 2461 GACTTGCTCATTTTTCGAGTTTAACTAGACTACACTATTTGAGTTTAAATTTTGTCC 2520
Db 2461 GACTTGCTCATTTTTCGAGTTTAACTAGACTACACTATTTGAGTTTAAATTTTGTCC 2520
QY 2521 TCTAGGATTTATTTCTGTGTCCAAAAAA 2563
Db 2521 TCTAGGATTTATTTCTGTGTCCAAAAAA 2563

```

RESULT 2

ABV23195 standard; cDNA; 3634 BP.

AC ABV23195;

DE Human prostate expression marker cDNA 23186.

Human prostate expression marker cDNA 23186.
 Human prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 pharmacogenomic marker; gene; ss.

OS Homo sapiens.
 XX
 XX
 PN WO200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US05171.
 XX

PR 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Schlegel R, Endege WO, Monahan JE;

WPI; 2001-662795/76.

Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer -

Claim 1; Page 4168-4169; 11750pp; English.

The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for:
 (a) assessing whether a patient is afflicted with prostate cancer;
 (b) monitoring the progression of prostate cancer in a patient;
 (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;
 (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;
 (e) selecting a composition for inhibiting prostate cancer in a patient;
 (f) assessing the prostate cell carcinogenic potential of a compound;
 (g) determining whether prostate cancer has metastasized in a patient;
 (h) assessing the aggressiveness or indolence of prostate cancer in a patient;
 (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

Sequence 3634 BP; 876 A; 966 C; 869 G; 918 T; 5 other;

Query Match 98.8%; Score 2531.6; DB 23; Length 3634;

Best local similarity 99.8%; Pred. No. 0;

Matches 2556; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

```

QY 1 GGGCGGGGAGGTCGAGAGCTGAGAGTCTGCTCTGTTCTCTGTGCGCACCGTGGCT 60
Db 71 GGGCGGGGAGGTCGAGAGCTGAGAGTCTGCTCTGTTCTCTGTGCGCGCGCTGGCT 130
QY 61 GTCGGGCTGCTGGGCTGCGGAACAGACAGAGGCGTGGGCGCACAGCACTCAGAAAGCGA 120
Db 131 GTCGGGCTGCTGGGCTGCGGAACAGACAGAGGCGTGGGCGCACAGCACTCAGAAAGCGA 190
QY 121 CGCAGCTGACGCGAGCGCGCGCAGAGAGGTCGAGAGGCGTGGGCGCGCGCGC 180
Db 191 CGCAGCTGACGCGAGCGCGCGCAGAGAGGTCGAGAGGCGTGGGCGCGCGCGC 250
QY 181 GGGCAGCGCGCGCGCGCGCAGAGAGGCGCGCGCGCGCGCGCTGAGCGCTGGCC 240
Db 251 GGGCAGCGCGCGCGCGCGCGCAGAGAGGCGCGCGCGCGCGCTGAGCGCTGGCC 310
QY 241 GGGCGGGGCGGGAATGTCATGCTGACGCGATGCCCTGCGCGCGGGGTGCGGAGGAGC 300
Db 311 GGGCGGGGCGGGAATGTCATGCTGACGCGATGCCCTGCGCGCGGGGTGCGGAGGAGC 370
QY 301 TGAAGCAGGCGCAAGAGATCGAGAGCGCGGAGAGTACTCTTCATGCGCACCGTACCA 360
Db 371 TGAAGCAGGCGCAAGAGATCGAGAGCGCGGAGAGTACTCTTCATGCGCACCGTACCA 430
QY 361 AGCGGCCCAAGAGAAATCCAGTTTGTGATGATGATGATGATGATGATGATGATGATGAT 420
Db 431 AGCGGCCCAAGAGAAATCCAGTTTGTGATGATGATGATGATGATGATGATGATGATGAT 490
QY 421 CCAAACTGCGCGCAAGATCTTGTCTGCTGCTGATCTCAGATCTCAGATCTCAGATCTCAG 480
Db 491 CCAAACTGCGCGCAAGATCTTGTCTGCTGCTGATCTCAGATCTCAGATCTCAGATCTCAG 550

```

```
QY 481 GTTCAGCTGCATCCTACACAGATAGCTCTGATGATGAGGTTTCTCCCGAGAGAAGCAGC 540
DB 551 GTTCAGCTGCATCCTACACAGATAGCTCTGATGATGAGGTTTCTCCCGAGAGAAGCAGC 610
QY 541 AAACCAACTCCAAAGGCGAGCAGCAATTTCTGTGTGAAGAATCAAGCAGGAGAGATTG 600
DB 611 AAACCAACTCCAAAGGCGAGCAGCAATTTCTGTGTGAAGAATCAAGCAGGAGAGATTG 670
QY 601 GACGCGGGAGATTGAGATTGACAGCAAGACATGTCTGCTCTGATTCTACTCAGGAAC 660
DB 671 GACGCGGGAGATTGAGATTGACAGCAAGACATGTCTGCTCTGATTCTACTCAGGAAC 730
QY 661 GTGCTCAGGGGAGAGAGCCCTTGCTGCTGCTAAATAGTGGCTGTACACATCAG 720
DB 731 GTGCTCAGGGGAGAGAGCCCTTGCTGCTGCTAAATAGTGGCTGTACACATCAG 790
QY 721 CCCAGACAGCGGTGTGATGAGACACTCTGTCCTGGGCTCAGTGCCTGCTGTG 780
DB 791 CCCAGACAGCGGTGTGATGAGACACTCTGTCCTGGGCTCAGTGCCTGCTGTG 850
QY 781 CTGTACATCTACTCAACTCAGATGAAGTAGTGCAGACACTGGCTGAGGCTGAGT 840
DB 851 CTGTACATCTACTCAACTCAGATGAAGTAGTGCAGACACTGGCTGAGGCTGAGT 910
QY 841 CAGTGTGCTTGAAGGCGAGTCAAGATGACTTCTGCTGTATTTGACCGCTGTG 900
DB 911 CAGTGTGCTTGAAGGCGAGTCAAGATGACTTCTGCTGTATTTGACCGCTGTG 970
QY 901 TGAACATGATGGGTGCGAGGCCAATGATCCTGATGATGGGAGACTTAACCCACT 960
DB 971 TGAACATGATGGGTGCGAGGCCAATGATCCTGATGATGGGAGACTTAACCCACT 1030
QY 961 GGGTTTATAGAAGTATCCAAACGTGTTTAAAGAAGTCCGAGGCTGTGTGGAAGAGCG 1020
DB 1031 GGGTTTATAGAAGTATCCAAACGTGTTTAAAGAAGTCCGAGGCTGTGTGGAAGAGCG 1090
QY 1021 TGACTGGTGTACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTGCGGCA 1080
DB 1091 TGACTGGTGTACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTGCGGCA 1150
QY 1081 TGAACGTCAATGATTTCTGTACCAACAAGAAGTTGATTAATCTGACTGCGGAGAT 1140
DB 1151 TGAACGTCAATGATTTCTGTACCAACAAGAAGTTGATTAATCTGACTGCGGAGAT 1210
QY 1141 CCATTTTGATGGCTGGAAGAGGACACAGATGTGATGTTGGGAAACAAGTGTG 1200
DB 1211 CCATTTTGATGGCTGGAAGAGGACACAGATGTGATGTTGGGAAACAAGTGTG 1270
QY 1201 TGTGTGCTATGCTGAGTAGGCAAGGCTGCTGCTCTCAAAAGCTCTTGAGCA 1260
DB 1271 TGTGTGCTATGCTGAGTAGGCAAGGCTGCTGCTCTCAAAAGCTCTTGAGCA 1330
QY 1261 TTGCTACATTAACCGAAATCGACCCCATCTGTGCTCTGCAAGGCTGATGAGT 1320
DB 1331 TTGCTACATTAACCGAAATCGACCCCATCTGTGCTCTGCAAGGCTGATGAGT 1390
QY 1321 GGGTGAAGCTAAATGAAGTCAATCCGCAAGTGTGATGCTAATACTTGACAGAA 1380
DB 1391 GGGTGAAGCTAAATGAAGTCAATCCGCAAGTGTGATGCTAATACTTGACAGAA 1450
QY 1381 ATAGAATGTAGTGAACGGGAGCACTTGATGCAATGAATAAGTGTATGATGCA 1440
DB 1451 ATAGAATGTAGTGAACGGGAGCACTTGATGCAATGAATAAGTGTATGATGCA 1510
QY 1441 ATATGGCCACTCCAAACAGAAATCGATGTGACAGCCCTCCGCACTCCGAGCTGACGT 1500
DB 1511 ATATGGCCACTCCAAACAGAAATCGATGTGACAGCCCTCCGCACTCCGAGCTGACGT 1570
QY 1501 GGGAGCGAGTACGTTCTCAGGTGACCATGTCTGCGCAGATGGCAAACGAGTGTCC 1560
DB 1571 GGGAGCGAGTACGTTCTCAGGTGACCATGTCTGCGCAGATGGCAAACGAGTGTCC 1630
QY 1561 TCCTGGCAGAGGCTCTCTACTCAATTTGAGCTGTCCACAGTTCACACCTTTGTTCTGT 1620

DB 1631 TCCTGGCAGAGGCTCTCTACTCAATTTGAGCTGTCCACAGTTCACACCTTTGTTCTGT 1690
QY 1621 CCATCAGAGCCACAACACAGGCTTTGGCACTGATGAAGTCTATATGACCCGAGGGC 1680
DB 1691 CCATCAGAGCCACAACACAGGCTTTGGCACTGATGAAGTCTATATGACCCGAGGGC 1750
QY 1681 GATCAAGCAGGATGTGACTTGTCTTCCTAAGAAATGATGATACGTTGCCAGCTTGC 1740
DB 1751 GATCAAGCAGGATGTGACTTGTCTTCCTAAGAAATGATGATACGTTGCCAGCTTGC 1810
QY 1741 ATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1800
DB 1811 ATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1870
QY 1801 GACTCAACAAAATGGGCCATTCAAACCTAATTATTACAGATACTAATGACCATACTAC 1860
DB 1871 GACTCAACAAAATGGGCCATTCAAACCTAATTATTACAGATACTAATGACCATACTAC 1930
QY 1861 CAAGACCAAGTCCACCTGAACACACACTCTAAAGAATATTTTAAAGATACTTTAT 1920
DB 1931 CAAGACCAAGTCCACCTGAACACACACTCTAAAGAATATTTTAAAGATACTTTAT 1990
QY 1921 TTCTCTTACTCCTTCTCTCTGATTTTCTCTATATTTTCTATCTCTGTTTTCATC 1980
DB 1991 TTCTCTTACTCCTTCTCTCTGATTTTCTCTATATTTTCTATCTCTGTTTTCATC 2050
QY 1981 TCATATCCAAAGTCTGACAGCCACACAGAACTGCTTCATGGCTCTTTAGATGAATA 2040
DB 2051 TCATATCCAAAGTCTGACAGCCACACAGAACTGCTTCATGGCTCTTTAGATGAATA 2110
QY 2041 GAAGTTCAAGGCTCCCTCACTCTAGTCACTAAAGAAGATTCTACTCCGCCAGCAGAAA 2100
DB 2111 GAAGTTCAAGGCTCCCTCACTCTAGTCACTAAAGAAGATTCTACTCCGCCAGCAGAAA 2170
QY 2101 GGTGATTTCTCTTTTACCATTTCTGGGCACTTTAGTCTTAATAGGTAATTAACA 2160
DB 2171 GGTGATTTCTT-TCTTACCATTTCTGGGCACTTTAGTCTTAATAGGTAATTAACA 2229
QY 2161 GAAATGCTAAGTACCTTCTCTGCAACAATCTGCAATGTCTAAATGCGCTTAAAGA 2220
DB 2230 GAAATGCTAAGTACCTTCTCTGCAACAATCTGCAATGTCTAAATGCGCTTAAAGA 2289
QY 2221 GCCCATTTCTTAGCTGTGAATCAGTCTCTTTCATCTCTTCAGAGAACAGGATG 2280
DB 2290 GCCCATTTCTTAGCTGTGAATCAGTCTCTTTCATCTCTTCAGAGAACAGGATG 2349
QY 2281 ACCTACCCGGCAGGTAGGTAGATGTGGGTGTGATGTTAATTTCCCTTAGAAGTCCA 2340
DB 2350 ACCTACCCGGCAGGTAGGTAGATGTGGGTGTGATGTTAATTTCCCTTAGAAGTCCA 2409
QY 2341 AGCCCTGTTCTGCGTAAGGTGATGTCCAGTTCAGAGATGTGATTAATGAGCATGG 2400
DB 2410 AGCCCTGTTCTGCGTAAGGTGATGTCCAGTTCAGAGATGTGATTAATGAGCATGG 2469
QY 2401 CTGTTAAGATCAGAGGCGCACTTGATTTATAGTATAGCCCTTCCCTCACCACCA 2460
DB 2470 CTGTTAAGATCAGAGGCGCACTTGATTTATAGTATAGCCCTTCCCTCACCACCA 2529
QY 2461 GACTGCTCATTTTTCGAGTTTAACTAGACTACACTCTAATTTGAGTTAATTTGTCC 2520
DB 2530 GACTGCTCATTTTTCGAGTTTAACTAGACTACACTCTA-TTGAGTTAATTTGTCC 2588
QY 2521 TCTAGATTATTTCTGTGTCCAAAAAATTTTAAAAA 2562
DB 2589 TCTAGATTATTTCTGTGTCCAAAAAATTTTAAAAA 2630

RESULT 3
ABV29032
ID ABV29032 standard; cdna; 3634 BP.
XX
AC ABV29032;
```

XX 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 29023.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-18319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer
XX
PS Claim 1; Page 6147-6148; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 3634 BP; 876 A; 966 C; 869 G; 918 T; 5 other;
Query Match 98.8%; Score 2531.6; DB 23; Length 3634;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2556; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 241 GGGCGGGGGGAAATGTCGATGCCCTGACGCGGATGCCCTGCCGGGGTCCGGGAGAGC 300
DB 311 GGGCGGGGGGAAATGTCGATGCCCTGACGCGGATGCCCTGCCGGGGTCCGGGAGAGC 370
QY 301 TGAAGCAGGCCAAGAGATCGAGGACGCCGAGAACTCTCTCATGCGCACCGTACCA 360
DB 371 TGAAGCAGGCCAAGAGATCGAGGACGCCGAGAACTCTCTCATGCGCACCGTACCA 430
QY 361 AGCGCCCAAGAAATCCAGTTTGTGATGACATGACGAGATTCCAAATTCGCCA 420
DB 431 AGCGCCCAAGAAATCCAGTTTGTGATGACATGACGAGATTCCAAATTCGCCA 490
QY 421 CCAAACTGGCCGAAGATCTTGTCTCGCTCGATCTCAGTCTCCACTGACAGCTACA 480
DB 491 CCAAACTGGCCGAAGATCTTGTCTCGCTCGATCTCAGTCTCCACTGACAGCTACA 550
QY 481 GTTCAGCTGATCTCTACACAGATAGCTCTGATGATGAGTTTCTCCCGAGAGAGCAGC 540
DB 551 GTTCAGCTGATCTCTACACAGATAGCTCTGATGATGAGTTTCTCCCGAGAGAGCAGC 610
QY 541 AAACCACTCCAAAGGAGCAGCAATTTCTGTGTGAAGAACTCAAGCAGGAGAAATTG 600
DB 611 AAACCACTCCAAAGGAGCAGCAATTTCTGTGTGAAGAACTCAAGCAGGAGAAATTG 670
QY 601 GACCGCGGAGATTGAGATTGACAGCAAGACATGTCGCTGATTTCACTCAGGAAC 660
DB 671 GACCGCGGAGATTGAGATTGACAGCAAGACATGTCGCTGATTTCACTCAGGAAC 730
QY 661 GTGCTCAGGGGAGAGAGCCCTTGCTGCTGCTAAATAGTGGGCTGTACACACATCACAG 720
DB 731 GTGCTCAGGGGAGAGAGCCCTTGCTGCTGCTAAATAGTGGGCTGTACACACATCACAG 790
QY 721 CCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGGGGGCTCAGTGCCGCTGTG 780
DB 791 CCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGGGGGCTCAGTGCCGCTGTG 850
QY 781 CTGTACATCTACTCACTCAGATGAAGTAGTGTGACAGCTGAGGCTGAGGTTG 840
DB 851 CTGTACATCTACTCACTCAGATGAAGTAGTGTGACAGCTGAGGCTGAGGTTG 910
QY 841 CAGTGTGCTTGAAGGGGAGCTGAGAGATGACTTCTGTGTGATGACCGCTGTG 900
DB 911 CAGTGTGCTTGAAGGGGAGCTGAGAGATGACTTCTGTGTGATGACCGCTGTG 970
QY 901 TGAACATGATGGGTGGAGGCCAATGATCTGATGATGGGAGACTTAACCCACT 960
DB 971 TGAACATGATGGGTGGAGGCCAATGATCTGATGATGGGAGACTTAACCCACT 1030
QY 961 GGGTTTAAAGAGTATCCAAACGTGTTTAAAGAGATCCGAGGCAATTGTGAAGAGCG 1020
DB 1031 GGGTTTAAAGAGTATCCAAACGTGTTTAAAGAGATCCGAGGCAATTGTGAAGAGCG 1090
QY 1021 TGAAGTGTGTTACAGAGGCTGTATCAGCTCTCCAAAGCTGGGAGCTGTGTCCGCCA 1080
DB 1091 TGAAGTGTGTTACAGAGGCTGTATCAGCTCTCCAAAGCTGGGAGCTGTGTCCGCCA 1150
QY 1081 TGAAGTCAATGATCTGTACCAACAGAGATTGATTAATCTGTACTGCTGCCGAGAT 1140
DB 1151 TGAAGTCAATGATCTGTACCAACAGAGATTGATTAATCTGTACTGCTGCCGAGAT 1210
QY 1141 CCATTTTGATGGCCTGAAGAGACACAGATGTGATGTTGTGGGAAACAAGTGTGG 1200
DB 1211 CCATTTTGATGGCCTGAAGAGACACAGATGTGATGTTGTGGGAAACAAGTGTGG 1270
QY 1201 TGTGTGCTATGTGAGGTAGGCAAGGGCTGTGTGCTGTCTCAAACTCTTGAGCAA 1260
DB 1271 TGTGTGCTATGTGAGGTAGGCAAGGGCTGTGTGCTGTCTCAAACTCTTGAGCAA 1330
QY 1261 TTGCTACATTACCGAAATGACCCCATCTGTGCTGACAGGCTGCATGATGGGTCA 1320
DB 1331 TTGCTACATTACCGAAATGACCCCATCTGTGCTGACAGGCTGCATGATGGGTCA 1390
QY 1321 GGGTGTAAAGCTAAATGAAGTCAATCCGGCAAGTGCATGTCTAATTACTTGACAGGAA 1380

```

Db 1391 GGGTGAAGCTAAATGAAGTCATCCGGCAAGTCGATGCTAATACTTGACACAGGAA 1450
OY 1381 ATAGAATGTAGTGAACGCGGAGCACTTGATCGCATGAAACAGTTGTATCGTATGCA 1440
Db 1451 ATAGAATGTAGTGAACGCGGAGCACTTGATCGCATGAAACAGTTGTATCGTATGCA 1510
OY 1441 ATATGGCCACTCCACACAGAAATCGATGTGACCAGCCTCCGCACTCCGAGCTGACGT 1500
Db 1511 ATATGGCCACTCCACACAGAAATCGATGTGACCAGCCTCCGCACTCCGAGCTGACGT 1570
OY 1501 GGGAGCGAGTACGTTCTCAGGTGAGCATGTCACTCTGGCCAGATGGCAAACGAGTTGCC 1560
Db 1571 GGGAGCGAGTACGTTCTCAGGTGAGCATGTCACTCTGGCCAGATGGCAAACGAGTTGCC 1630
OY 1561 TCCTGGCAGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTTCCACCTTTGTTCTGT 1620
Db 1631 TCCTGGCAGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTTCCACCTTTGTTCTGT 1690
OY 1621 CCATCACAGCCACACACAGGCTTTGGCAGTGTAGAACTCTATAATGACACCCGAGGGGC 1680
Db 1691 CCATCACAGCCACACACAGGCTTTGGCAGTGTAGAACTCTATAATGACACCCGAGGGGC 1750
OY 1681 GATACAGCAGAGTGTGTACTTGTCTTCTTGAAGAAATGATGAATACGTTGCCAGCTTGC 1740
Db 1751 GATACAGCAGAGTGTGTACTTGTCTTCTTGAAGAAATGATGAATACGTTGCCAGCTTGC 1810
OY 1741 ATCTGCCATCATTTGATGCCCACTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1800
Db 1811 ATCTGCCATCATTTGATGCCCACTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1870
OY 1801 GACTCAACAAAAATGGGCCATTCAAACCTAATTATACAGATACTAATGACCATACTAC 1860
Db 1871 GACTCAACAAAAATGGGCCATTCAAACCTAATTATACAGATACTAATGACCATACTAC 1930
OY 1861 CAAGACCAAGTCCACTGAAACACACACACTCTAAGAAATATTTTAAAGATACTTTAT 1920
Db 1931 CAAGACCAAGTCCACTGAAACACACACACTCTAAGAAATATTTTAAAGATACTTTAT 1990
OY 1921 TTTCTTCTTACTCTTCTCTCTCTGATTTTTTCCATAATTTCAATCTTGTGTTTTCATC 1980
Db 1991 TTTCTTCTTACTCTCTCTCTCTGATTTTTTCCATAATTTCAATCTTGTGTTTTCATC 2050
OY 1981 TCATTATCCAAAGTCTGACAGACACACAGAACTGCTTCATGCGCTTTAGATGAATA 2040
Db 2051 TCATTATCCAAAGTCTGACAGACACACAGAACTGCTTCATGCGCTTTAGATGAATA 2110
OY 2041 GAAGTTCAGGTCCTCACTCTAGTCACTAAGAAAGATTTTACTCCCCAGCCAGAAA 2100
Db 2111 GAAGTTCAGGTCCTCACTCTAGTCACTAAGAAAGATTTTACTCTCCAGCCAGAAA 2170
OY 2101 GGTGATCTCTCTTACCATTTCTGGGCACTTGTCTTAATTAGGTACCTTATTAACA 2160
Db 2171 GGTGATCTCTTCTTACCATTTCTGGGCACTTGTCTTAATTAGGTACCTTATTAACA 2229
OY 2161 GGAATGCTAAGGTACCTTCTGTGGAACAATCTGCAATGTCTAAATCGCCTTAAAGA 2220
Db 2230 GGAATGCTAAGGTACCTTCTGTGGAACAATCTGCAATGTCTAAATCGCCTTAAAGA 2289
OY 2221 GCCCATTTCTAGCTGCTGAAATCAGTCTCTTCACTTCTCAGAGAAGCAGGATGCT 2280
Db 2290 GCCCATTTCTAGCTGCTGAAATCAGTCTCTTCACTTCTCAGAGAAGCAGGATGCT 2349
OY 2281 ACCTACCCGCGAGTAGGTAGATGTGGGTGTCATGTTAATTTCCCTTAGAAGTTCCA 2340
Db 2350 ACCTACCCGCGAGTAGGTAGATGTGGGTGTCATGTTAATTTCCCTTAGAAGTTCCA 2409
OY 2341 AGCCCTGTTCTCGCTAAAGGTGATGTCCAGTTCCAGATGTGTATATGAGCATGG 2400
Db 2410 AGCCCTGTTCTCGCTAAAGGTGATGTCCAGTTCCAGATGTGTATATGAGCATGG 2469
OY 2401 CTGTTAAGATCAGGAGCCCACTGGATTTATAGTATAGCCCTTCTCCACTCCACCA 2460

```

```

Db 2470 CTGTTAAGATCAGGAGCCCACTTGATTTATAGTATAGCCCTTCTCCACTCCACCA 2529
OY 2461 GACTGCTCATTTTTCGAGTTTAACTAGACTACACTCTAATTGAGTTAATTTGTCC 2520
Db 2530 GACTGCTCATTTTTCGAGTTTAACTAGACTACACTCTA-TTGAATTAAATTTGTCC 2588
OY 2521 TCTAGATTATTTCTGTGTCCAAAAA 2562
Db 2589 TCTAGATTATTTCTGTGTCCAAAAA 2630

RESULT 4
AAV73924
ID AAV73924 standard; DNA; 3616 BP.
XX
AC AAV73924;
XX
DT 04-MAR-1999 (first entry)
XX
DE Human SAHH DNA #1.
XX
KW S-adenosyl-5-homocysteine hydrolase; SAHH; human; drug screening;
KW treatment; infection; cancer; autoimmune disease; detection; diagnosis;
KW gene mapping; antisense; therapy; antagonist; immunoassay; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 56..1558
FT /*tag= a
FT /product= "SAHH"
XX
PN US5854023-A*
XX
PD 29-DEC-1998.
XX
PF 17-JUL-1997; 97US-0896005.
XX
PR 17-JUL-1997; 97US-0896005.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Corley NC; Hillman UT; Lal P, Shah P;
XX
DR WPI; 1999-094906/08.
DR P-PSDB; AAM90061.
XX
PT Nucleic acid encoding human S-adenosyl-5-homocysteine hydrolase -
PT for production of recombinant enzyme, useful for diagnosis,
PT treatment and prevention of cancers, infections and autoimmune
PT diseases
XX
PS Disclosure; Fig 1A-1; 40pp; English.
XX
CC This sequence encodes a human S-adenosyl-5-homocysteine hydrolase (SAHH).
CC The SAHH protein can be used to generate specific antibodies and in drug
CC screening to identify specific binding agents. Antagonists of the
CC protein are used to treat or prevent a wide range of viral, bacterial,
CC fungal, parasitic, protozoal or helminthic infections, many cancers
CC (leukemia, lymphoma or solid tumours), and many autoimmune diseases
CC (e.g. acquired immune deficiency syndrome, allergy, asthma, diabetes
CC mellitus, multiple sclerosis etc). All these conditions may be treated by
CC expressing antisense sequences, triplex-forming agents or ribozymes
CC directed against the nucleic acid. The nucleic acid and its fragments can
CC be used as probes or primers for detecting and quantifying gene
CC expression, for diagnosis or monitoring of disease, to identify genetic
CC variations, mutations or polymorphisms, in gene mapping and as antisense
CC therapeutics. Antibodies are used directly as antagonists, indirectly to
CC deliver active agents to SAHH-expressing cells, to diagnose and monitor
CC diseases in standard immunoassays, in competitive drug screens and to
CC isolate the protein from natural sources.
XX
SQ Sequence 3616 BP; 1017 A; 782 C; 826 G; 991 T; 0 other;

```

Query Match 83.0%; Score 2126.8; DB 20; Length 3616;
Best Local Similarity 97.4%; Pred. No. 0;
Matches 2214; Conservative 0; Mismatches 32; Indels 28; Gaps 4;

```
QY 290 CCGGAGAGAGCTGAAGCAGGCCAAGAGATCGAGGACGCCGAGAGTACTCTTCATGCG 349
Db 1 CAGGAGAGAGCTGAAGCAGGCCAAGAGATCGAGGACGCCGAGAGTACTCTTCATGCG 60
QY 350 CACCGTCACCAAGCGGCCAAGAGCAAAATCCAGTTTGTGATGACATGAGAGTTCAC 409
Db 61 CACCGTCACCAAGCGGCCAAGAGCAAAATCCAGTTTGTGATGACATGAGAGTTCAC 120
QY 410 CAAATTCGCCCAAAACTGGCCGAGATCTTGTCTGCTCGATCTCAGAGTCTCCAC 469
Db 121 CAAATTCGCCCAAAACTGGCCGAGATCTTGTCTGCTCGATCTCAGAGTCTCCAC 180
QY 470 TGACAGCTACAGTTCAGCTGCATCTTACACAGATAGCTCTGATGATGAGGTTCTCCCG 529
Db 181 TGACAGCTACAGTTCAGCTGCATCTTACACAGATAGCTCTGATGATGAGGTTCTCCCG 240
QY 530 AGAGAGCAGCAAAACCACTCCAGGGGAGCAGCAATTTCTGTGAGAGCAATCAAGCA 589
Db 241 AGAGAGCAGCAAAACCACTCCAGGGGAGCAGCAATTTCTGTGAGAGCAATCAAGCA 300
QY 590 GCGAGAATTTGAGCGCCGAGATGAGATTGACAGCAAGACATGTCGTCTGATTTTC 649
Db 301 GCGAGAATTTGAGCGCCGAGATGAGATTGACAGCAAGACATGTCGTCTGATTTTC 360
QY 650 ACTCAGGAACAGTGTCTCAGGGGGAGAGCCCTTGCTGCTAAATAGTGGCTGTAC 709
Db 361 ACTCAGGAACAGTGTCTCAGGGGGAGAGCCCTTGCTGCTAAATAGTGGCTGTAC 420
QY 710 ACACATCACAGCCCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGGGGCTCAGTG 769
Db 421 ACACATCACAGCCCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGGGGCTCAGTG 480
QY 770 CCGCTGCTGTCTGTATACATCTACTCACTCAGATGAGTAGTGCAGCAGCTGCTGA 829
Db 481 CCGCTGCTGTCTGTATACATCTACTCACTCAGATGAGTAGTGCAGCAGCTGCTGA 540
QY 830 GCGTGGAGTTGACGTGTTGCTTGAAGGGGAGTCAAGATGACTTCTGTGTTAT 889
Db 541 GCGTGGAGTTGACGTGTTGCTTGAAGGGGAGTCAAGATGACTTCTGTGTTAT 600
QY 890 TGAACGCTGTGTGACATGATGGGTGACAGGCCAATCATGATCCTGATGATGGGAGA 949
Db 601 TGAACGCTGTGTGACATGATGGGTGACAGGCCAATCATGATCCTGATGATGGGAGA 660
QY 950 CTTAACCCACTGGGTTTATAGAAGTATCCAAAGTGTTAAGAAGATCCGAGGCAATTGT 1009
Db 661 CTTAACCCACTGGGTTTATAGAAGTATCCAAAGTGTTAAGAAGATCCGAGGCAATTGT 720
QY 1010 GGAAGAGAGCTGAGTGTGTTACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTG 1069
Db 721 GGAAGAGAGCTGAGTGTGTTACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTG 780
QY 1070 TGTTCGGCCATGAAGCTCATGATCTGTACCAAAAGAGATTGATACTGTACTG 1129
Db 781 TGTTCGGCCATGAAGCTCATGATCTGTACCAAAAGAGATTGATACTGTACTG 840
QY 1130 CTGCGAAGATTCATTTTGGATGGCCTGAAGAGACCAAGATGTGTTGTGGAA 1189
Db 841 CTGCGAAGATTCATTTTGGATGGCCTGAAGAGACCAAGATGTGTTGTGGAA 900
QY 1190 ACAAGTGTGTGTGTGCTATGTTGAGGTAGGCAAGGCTGCTGTGCTCTCAAGC 1249
Db 901 ACAAGTGTGTGTGTGCTATGTTGAGGTAGGCAAGGCTGCTGTGCTCTCAAGC 960
QY 1250 TCTTGAGCAATGTCTACATTACGAAATCGACCCATCTGTCTGAGGCGCTGCAT 1309
Db 961 TCTTGAGCAATGTCTACATTACGAAATCGACCCATCTGTCTGAGGCGCTGCAT 1020
```

```
QY 1310 GGATGGTTACAGGCTGCTAAAGCTAAATGAGTTCATCCGGCAAGTGCATGCTGTAATAC 1369
Db 1021 GGATGGTTACAGGCTGCTAAAGCTAAATGAGTTCATCCGGCAAGTGCATGCTGTAATAC 1080
QY 1370 TPGCACAGGAATAAGATGTAGTACACAGGAGCACTTGATCGCATGAACAAAGTGTG 1429
Db 1081 TTGCACAGGAATAAGATGTAGTACACAGGAGCACTTGATCGCATGAACAAAGTGTG 1140
QY 1430 TATCGTATGCAATATGGGCCACTCCAAACACAGAAATCGATGTGACCAGCCTCCGACTCC 1489
Db 1141 TATCGTATGCAATATGGGCCACTCCAAACACAGAAATCGATGTGACCAGCCTCCGACTCC 1200
QY 1490 GGAGCTGACGTGGAGCGAGTACGTTCTCAGGTGACCATGTCATCTGGCCAGATGGCA 1549
Db 1201 GGAGCTGACGTGGAGCGAGTACGTTCTCAGGTGACCATGTCATCTGGCCAGATGGCA 1260
QY 1550 ACGAGTGTCTCTCTGCGCAGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTCCAC 1609
Db 1261 ACGAGTGTCTCTCTGCGCAGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTCCAC 1320
QY 1610 CTTGTTCTGTCATCACAGCCACACACAGGCTTTGGCACTGATAGAACTCTATATGC 1669
Db 1321 CTTGTTCTGTCATCACAGCCACACACAGGCTTTGGCACTGATAGAACTCTATATGC 1380
QY 1670 ACCGAGGGGCAATACAGAGAGTGTACTTGTCTTCTTAAGAAATGATGAATACGT 1729
Db 1381 ACCGAGGGGCAATACAGAGAGTGTACTTGTCTTCTTAAGAAATGATGAATACGT 1440
QY 1730 TGCCAGCTGCTATGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGC 1789
Db 1441 TGCCAGCTGCTATGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGC 1500
QY 1790 AAAATATCTGGAGCTCAACAAAATGGGCCATTCAAACCTAATTTACAGATACTAATG 1849
Db 1501 AAAATATCTGGAGCTCAACAAAATGGGCCATTCAAACCTAATTTACAGATACTAATG 1560
QY 1850 GACCATACTACCAAGACCACTCCACCTGAACACACACTCTAAAGAAATATTTTAAAG 1909
Db 1561 GACCATACTACCAAG-----CAATCAATCGGAGTCTTTGGG 1597
QY 1910 ATAATTTATTTCTCTACTCCTCTCTGATTTTCTCTATATTTCTATCTTCT 1969
Db 1598 CCATGCTGCCAGTCCACCTGA---CTTCTCTGATTTTCTCTATATTTCTATCTTCT 1654
QY 1970 GTTTTTCATCTCATTAATCCAAAGTTCGACAGACACACAGAACTTGCTCATGGCTCT 2029
Db 1655 GTTTTTCATCTCATTAATCCAAAGTTCGACAGACACACAGAACTTGCTCATGGCTCT 1714
QY 2030 TAGATGAATAGAAAGTCAAGGCTCCCTCACTCTAGTCACTAAGAAAGATTTTACTCCCC 2089
Db 1715 TAGATGAATAGAAAGTCAAGGCTCCCTCACTCTAGTCACTAAGAAAGATTTTACTCCCC 1774
QY 2090 CAGCCAGAAAGGTGATCTCTCTTTACCATTTCTGGGACTTTAGCTTAATTAGGTA 2149
Db 1775 CAGCCAGAAAGGTGATCTT-TCTTACCATTTCTGGGACTTTAGCTTAATTAGGTA 1833
QY 2150 CCTTATTACAGGAATGCTAAGTACCTCTCTGTGAACAATCTGCAATGTCTAATC 2209
Db 1834 CCTTATTACAGGAATGCTAAGTACCTCTCTGTGAACAATCTGCAATGTCTAATC 1893
QY 2210 GCCTTAAAGAGCCCATTTCTTAGCTGCTGAATCAGTGTCTTTCACTTCTCAGAGAA 2269
Db 1894 GCCTTAAAGAGCCCATTTCTTAGCTGCTGAATCAGTGTCTTTCACTTCTCAGAGAA 1953
QY 2270 GCAGGATGTACCTACCGGAGAGGTAGGTAGATGTGGGTGTCATGTTAATTTCCCT 2329
Db 1954 GCAGGATGTACCTACCGGAGAGGTAGGTAGATGTGGGTGTCATGTTAATTTCCCT 2013
QY 2330 TAGAAGTTCCAAGCCCTGTTCTGCGTAAGGTGATGTCAGATGATGTAT 2389
Db 2014 TAGAAGTTCCAAGCCCTGTTCTGCGTAAGGTGATGTCAGATGATGTAT 2073
QY 2390 AATGAGCATGCTGTTAAGATCAGAGGCCCACTTGATTTATAGTATAGCCCTTCCCTC 2449
```


QY	1166	CACAGATGATGTTTGGTGGGAAACAAGTGGTGGTGGGCTATGGTGAGGTAGGCCAA	1225
Db	3852	AACAGACATGATGTTGGTGGAAAGCAAGTGGTGGTGGCTATGGAGAGGTGGGAA	3793
QY	1226	GGGCTGCTGTGCTGCTCTCAAGGCTCTTGAGACAAATGCTACATTACCGAAATCGACCC	1285
Db	3792	AGGGTCTGTGCTGCGCCCTGAAAGCCATGGGCTCCATGTGTATGTAACTGAAATTGACCC	3733
QY	1286	CATCTGTGCTCTGCAAGGCTGCATGGATGGGTTCAGGGTGGTAAAGCTAAATGAATCAT	1345
Db	3732	CATCTGTGCTCTGCAAGGCTGTATGGATGGATTTGACTGGTGAATTAATGAGTCAAT	3673
QY	1346	CCGGCAAG-TCGATGTCGTAATACCTTGACACAGAAATAAGATAGTGACACGGGAGC	1404
Db	3672	CCGACAAGTTGGACAATTGTATTACCTGTACAGGTAAACAAGAAATGTGGTAAACAGAGAGC	3613
QY	1405	ACTTGATCGCATGAAAAAACAGTTGTATCGTATGCAATATGGGCCACTCCAACACAGAAA	1464
Db	3612	ACTTGACCGTATGAAGAATAGCTGCATCGTTTGTAAACATGGGACATTCCAACACAGAGA	3553
QY	1465	TCGATGTGACCAGCCCTCCGCACTCC-GGAGCTGACGTGGAGCGAGTAACGTTCTCAGGTG	1523
Db	3552	TTGACGTGGCGAGTCTGCGGACACCAAGAACTGACTGGGAGCGAGTGAATCTCAAGTT	3493
QY	1524	GACCATGTCACTCGGCCAGATGGCAACAGAGTGTCTCTCGGACAGGCTCGTCACTC	1583
Db	3492	GACCATGTGATATGGCCTGATGGCAAGAGATAGTACTGCTGGCAGAGGCGCCCTGCTG	3433
QY	1584	AATTTGAGCTGCTCCACAGAGTTCACACCCTTGTCTGTCCATCACAGCCACAACAC-AGGC	1642
Db	3432	AACCTTAGCTGCTCCACAGTGCCTACATTGTGTCTCAATCACTGTACTACTCAAGGC	3373
QY	1643	TTTGGCACTGATAGAACTCTATATGCACCCGAGGGCGATACAGAGAGATGTGTACTT	1702
Db	3372	TCCTGCTTGTAGAGCTTTACAATGCTCCTGAGGGTGGCTATTAAGCAGAGATGCTCACT	3313
QY	1703	GCTTCCTAAGAAATGGATGAATACGTTGCCAGCTT-GCATCTGCCATCATTTGATGCC	1761
Db	3312	GTTGCCCAAGAGATGATGATGATATGTGGCCAGCTTACCACCTGCTTACCTTTGATGCC	3253
QY	1762	ACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAACAATAATGGGCCAT	1821
Db	3252	ACTTGACAGAGCTGACAGATGAACAAGCCCAAGTATCTGGGACTCAATTAAGATGGGCCCT	3193
QY	1822	TCAAACCTAATTATTACAGATACTAATGGACCATACTACCAAGGACCAGT	1871
Db	3192	TCAAGCTAATTACTACAGGTATTAAGTCTGTAACTCAAACCAAGATTT	3143

Accession Number	Source	Accession Number	Source
AAV73925	AAV73925 standard; DNA; 2226 BP.	AAV73925	AAV73925
AC	AAV73925;	AC	AAV73925;
DT	04-MAR-1999 (first entry)	DT	04-MAR-1999 (first entry)
DE	Human SAHH DNA #2.	DE	Human SAHH DNA #2.
KW	S-adenosyl-5-homocysteine hydrolase; SAHH; human; drug screening; treatment; infection; cancer; autoimmune disease; detection; diagnosis; gene mapping; antisense; therapy; antagonist; immunoassay; ss.	KW	S-adenosyl-5-homocysteine hydrolase; SAHH; human; drug screening; treatment; infection; cancer; autoimmune disease; detection; diagnosis; gene mapping; antisense; therapy; antagonist; immunoassay; ss.
OS	Homo sapiens.	OS	Homo sapiens.
PN	US5854023-A.	PN	US5854023-A.
PD	29-DEC-1998.	PD	29-DEC-1998.
PF	17-JUL-1997; 97US-0896005.	PF	17-JUL-1997; 97US-0896005.
PR	17-JUL-1997; 97US-0896005.	PR	17-JUL-1997; 97US-0896005.

PA	(INCY-) INCYTE PHARM INC.	
XX		
PI	Corley NC, Hillman JL, Lai P, Shah P;	
XX		
DR	WPI; 1999-094906/08.	
P	P-PSDB; AAW90061.	
XX		
PT	Nucleic acid encoding human S-adenosyl-5-homocysteine hydrolase -	
PT	for production of recombinant enzyme, useful for diagnosis,	
PT	treatment and prevention of cancers, infections and autoimmune	
PT	diseases	
XX		
PS	Claim 5; Column 39-42; 40pp; English.	
XX		
CC	This sequence encodes a human S-adenosyl-5-homocysteine hydrolase (SAHH).	
CC	The SAHH protein can be used to generate specific antibodies and in drug	
CC	screening to identify specific binding agents. Antagonists of the	
CC	protein are used to treat or prevent a wide range of viral, bacterial,	
CC	fungal, parasitic, protozoal or helminthic infections, many cancers	
CC	(leukaemia, lymphoma or solid tumours), and many autoimmune diseases	
CC	(e.g. acquired immune deficiency syndrome, allergy, asthma, diabetes	
CC	mellitus, multiple sclerosis etc). All these conditions may be treated by	
CC	expressing antisense sequences, triplex-forming agents or ribozymes	
CC	directed against the nucleic acid. The nucleic acid and its fragments can	
CC	be used as probes or primers for detecting and quantifying gene	
CC	expression, for diagnosis or monitoring of disease, to identify genetic	
CC	variations, mutations or polymorphisms, in gene mapping and as antisense	
CC	therapeutics. Antibodies are used directly as antagonists, indirectly to	
CC	deliver active agents to SAHH-expressing cells, to diagnose and monitor	
CC	diseases in standard immunoassays, in competitive drug screens and to	
CC	isolate the protein from natural sources.	
XX		
SQ	Sequence 2226 BP; 649 A; 455 C; 435 G; 675 T; 12 other;	
Query Match	31.5%; Score 806.2; DB 20; Length 2226;	
Best Local Similarity	99.3%; Pred. No. 1.6e-198;	
Matches 830; Conservative	0; Mismatches 4; Indels 2; Gaps 2;	
OY	1728 GTTGCCAGCTGCATCTGCCATCATTTGATGGCCACCTTACAGAGCTGACAGATGACC	1787
Db	1 GTTGCCAGCTGCATCTGCCATCATTTGATGGCCACCTTACAGAGCTGACAGATGACC	60
OY	1788 GCAAAATATCTGGGACTCAACAAAAATGGGCCATTCAAACCTAATTATTACAGATACTAA	1847
Db	61 GCAAAATATCTGGGACTCAACAAAAATGGGCCATTCAAACCTAATTATTACAGATACTAA	120
OY	1848 TGGACCACTACTACCAAGACAGCAGTCACCTGAACACACACTCTAAGAAGAAATATTTT	1907
Db	121 TGGACCACTACTACCAAGACAGCAGTCACCTGAACACACACTCTAAGAAGAAATATTTT	180
OY	1908 AGATAACTTTTATTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	1967
Db	181 AGATAACTTTTATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	240
OY	1968 TTGTTTTTCATCTCATTTATCCAGTCTTGCAGACACACAGGAACCTTGCTTCATGGCTC	2027
Db	241 TTGTTTTTCATCTCATTTATCCAAAGTTTGCAGACACACAGGAACCTTGCTTCATGGCTC	300
OY	2028 TTTAGATGAATAGAAAGTTGAGGCTCCCTCACCTAGTCACTAAGAAGGATTTTACTCC	2087
Db	301 TTTAGATGAATAGAAAGTTGAGGCTCCCTCACCTAGTCACTAAGAAGGATTTTACTCT	360
OY	2088 CCCAGCCCAGAAAGGTGATTTCTCTTACCAATTTCTGGGACTTAGTCTTAATTAAGG	2147
Db	361 CCCAGCCCAGAAAGGTGATTTCT-TCTTTACCAATTTCTGGGACTTAGTCTTAATTAAGG	419
OY	2148 TACCTTATTAACAGAAATGCTAAGGTACCTTCTCTGTGGAACAATCTGCAATGTCTAAA	2207
Db	420 TACCTTATTAACAGAAATGCTAAGGTACCTTCTCTGTGGAACAATCTGCAATGTCTAAA	479
OY	2208 TCGCCCTTAAAGAGCCCATTTCTTAGCTGCTGAAATCAGTGCTCTTTCACCTTCTTCAGAG	2267
Db	480 TCGCCCTTAAAGAGCCCATTTCTTAGCTGCTGAAATCAGTGCTCTTTCACCTTCTTCAGAG	539

KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li FWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions
XX
PS Claim 1; SEQ ID NO 40834; 21pp + sequence listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB16175), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (AB57737-AB572072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1776 BP; 437 A; 432 C; 479 G; 428 T; 0 other;
Query Match 25.3%; Score 649.2; DB 23; Length 1776;
Best Local Similarity 67.0%; Pred. No. 7.8e-158;
Matches 937; Conservative 0; Mismatches 458; Indels 3; Gaps 1;
QY 445 CTCGCTGATCTACAGTCTCTCCACTGACAGCTACAGTTCAGTGCATCTTACACAGATA 504
DB 275 CTCGGCGTTTCTGATGTCACCTGCTCATTCAGTTCACCTGTTTCCACCGAAGCT 334
QY 505 GCTCTGATGATGAGTTTCTCCCGAGAGAGACGACCAACTCCAGGGCAGCAGCA 564
DB 335 CCGACGAGAGACGTCCTCCCGAAGGACACCATCAGCGCACTCCCGCGGAGCACTG 394
QY 565 ATTCTGTGTGAAGACATCAAGCAGGAGATTTGGACCGCGGAGATGAGATTGACAG 624
DB 395 ACTTTGCGTGAAGACATCTCGAAGAGTGCATTGGAAGCGGGAGATCGAGATCGCCG 454
QY 625 AGCAAGACATGCTGCTGATTTCACTCAGGAACGTCGTCAGGGGAGAGAGCCCTTG 684
DB 455 AGTCGAGATGCCGGCATCATGACTCTGAGAAAGAGCGAAGATGAGAAAGCCCTTA 514
QY 685 CTGTGCTAAATAGTGGGCTGTACACACATCACAGCCAGACGCGTGTGATTGAGA 744
DB 515 AGGTGCCAATATCGTCGATGACACCCAGTCATGTCAGTCGCGCAGTGTGATGAGA 574
QY 745 CACTCTGTGCCCTGGGGCTCAGTGCCTGCTGTGTAACATCTACTCACTCAGA 804
DB 575 CCCTCGTCCAACCTGGGGCCACAGTTCGCTGGGCTGCTGCAACATTTATCCACAAA 634
QY 805 ATGAAGTAGCTGACAGACTGGCTGAGGCTGAGTGCAGTTCGCTTGAAGGGCAGT 864
DB 635 ACGCAGTTCGCCCGCTGTGCGACAGAGGGGGAATTCGATCTTCGCTGGCGGAGAGA 694
QY 865 CAGAAGATGACTTCTGTGTGTATGACCGCTGTGTGAACATGATGGGTGGCAGGCCA 924

DB 695 CCGAGAGAGAGTCTGTGTGTGCTTGACAGGGCCATCTACTCCGACGGCTGGCAGCCGA 754
QY 925 ACATGATCCTGGATGATGGGGAGACTTAACCCACTGGGTTTATAGAAGTATCCAAAG 984
DB 755 ACCTAATCCTGGACGACGGCGCGGATGCCACCGCACTTATGCTCAGAAGTACCCGACT 814
QY 985 TGTTTAAGAAGATCCGAGGCAATGTGAAGAGAGCGGTACTGTGTTCACAGGCTGTATC 1044
DB 815 ACTTCAAGGCCATTCGGGGCATCGTGAGAGAAAGTGTACCGGGGTGACCGGCTGTACA 874
QY 1045 AGCTCTCCAAAGCTGGGAAGCTCTGTGTCGGCCATGACGTCAATGATTCGTACCA 1104
DB 875 TGCTGTCAAAAGGGCGGAAACTTACTGTCCGGCCATCAACGTTAACGACTCAGTACCA 934
QY 1105 AACAGAGTTTGATACTGTACTGCTGCCGAGAATCCATTTTGGATGGCTGAAGAGGA 1164
DB 935 AGACAAAGTTTGATCTTCTACACGTCGTGACTCCATCCCTGACAGTCTGAACGCA 994
QY 1165 CCACAGATGTGATGTTTGGTGGGAACAAAGTGTGTGTGCTATGTGTAGGTAGGCA 1224
DB 995 CCACGGATATATATGTTTGGCGGAAGACAGTGTGATCTGTGGGTACGCTGATGTGGAA 1054
QY 1225 AGGCTGCTGTGCTGCTCTCAAAAGCTCTTGAGCAATTTGCTACATTACCGAAATGACCC 1284
DB 1055 AGGCTGTGCCAGTCTCCCTGAAGGGCCAAAGATGCAATGTTATGTTACGGAAGTGATC 1114
QY 1285 CCATCTGTGCTCTGACAGGCTGACGATGAGTGGGTGACGGGTGAAGCTTAATGAAGTCA 1344
DB 1115 CCATATGTGCTCTACAAAGCTGCATGATGATTCGCGGTGTACGGCTCAACGAGTCA 1174
QY 1345 TCCGCAAGTCGATGTCGTAATACTTGACAGAGAAATGAAGATGATGACACGGAGGC 1404
DB 1175 TCAGAGACGGTGTGATGTGTGTACGGCACTGGAACAAATGTTATTTACAGGATC 1234
QY 1405 ACTTGATCGCATGAAAAACAGTTGTATGCAATATGAGGACCTCCACACAGAGAA 1464
DB 1235 ACATGAATCGCATGAAGAAATGTTGTATTCCTGCAATATGGACATTCCTGCTCGGAGA 1294
QY 1465 TCGATGTGACAGCCTCCGACCTCCGAGCTGAGCTGGAGCGAGTACGTTCTCAGGTGG 1524
DB 1295 TTGATGTGAATGGCTTGCAATACCCCGAGGTACAGTGGAGCGTGTCCGTTCAAGTGG 1354
QY 1525 ACCATGTCATCTGGCCAGATGGCAACGAGTTGCTCTCTGCGCAGAGGGTCTCTACTCA 1584
DB 1355 ATCACATCAGGTGGCCGAGCGGACGAGATGATCATTTTGTCTCGCCGAGGAGACTGTGA 1414
QY 1585 ATTTAGCTGCTCCACAGTTCCACCTTTGTTCTGTCTCATCACAGCCACACAGGCTT 1644
DB 1415 ATTTGCTCTGTCCACCATTTGCTTGTGTATCCGTGCGCTCATCCACGAGCTT 1474
QY 1645 TGGCCTGATAGAACTCTATAATGACCCGAGGGCGATACAGCAGAGATGTACTTGC 1704
DB 1475 TGGCCCTGATTAAGTCTTCTCAGCGCC--AGGAAGATATAAGTCGATGTCTACTGC 1531
QY 1705 TTCTTAAGAAATGATGAATACGTTGCCAGCTTGCAATCTGCCATCATTTGATGCCACC 1764
DB 1532 TGCCAAAGAAATGATGATGATACGTGGCCAGTTGCAATCTGCCACCTTCGATGCTATC 1591
QY 1765 TTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAACAAAATGGGCAATCA 1824
DB 1592 TCACGGAAGCTCAGGATGAGCAGATCCAAAGTTATGGGCTTAACAAAGGCCGCTTTTA 1651
QY 1825 AACCTAATTATTACAGAT 1842
DB 1652 AAGCCAATTACTACAGGT 1669
RESULT 10
ABL79045
ID ABL79045 standard; cDNA; 553 BP.
XX
AC ABL79045;
XX

DT 17-MAY-2002 (first entry)
XX
DE Human ovarian cancer related CDNA clone SEQ ID NO:2023.
XX
KM Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200192581-A2.
XX
PD 06-DEC-2001.
XX
PF 29-MAY-2001; 2001WO-US17756.
XX
PR 26-MAY-2000; 2000US-207484P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Algate PA, Harlocker SL, Jones R;
XX
DR WPI; 2002-122075/16.
XX
PT Composition for therapy and diagnosis of ovarian cancer comprising
PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
PT polypeptide, antibody specific to polypeptide or T cell expressing
PT polypeptide
XX
PS Claim 1; SEQ ID 2023; 489pp; English.
XX
CC The present invention describes a composition (I) comprising: carriers
CC and immunostimulants; and a polypeptide (II) of a ovarian tumour
CC polypeptide encoded by a polynucleotide (III) having a CDNA sequence
CC (S1) from the 10912 nucleotide sequences as given in ABL77023 to
CC ABL87934, (III) encoding (II) having a sequence (S2), a T cell
CC population of (II), or antigen presenting cells that express (II).
CC (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to
CC (S1) can be used for detecting ovarian cancer in a patient's biological
CC sample preferably serum or ovarian tissue. The method comprises
CC contacting a biological sample from a patient with (IV), detecting the
CC amount of polynucleotide hybridising to (IV) and comparing the amount to
CC a predetermined cutoff value and thereby detecting ovarian cancer in the
CC patient, where the amount of polynucleotide hybridising to (IV) is
CC detected preferably by polymerase chain reaction (PCR). (I) comprising
CC (III) and/or (II) is useful for stimulating and/or expanding T cells
CC specific for an ovarian tumour protein comprising contacting T cells
CC with (III) or (II). (III) is useful in design and preparation of
CC ribozyme molecules for inhibiting expression of the tumour polypeptides
CC and proteins in tumour cells; and to isolate a full length gene from a
CC suitable library e.g., a tumour CDNA library using well known
CC techniques.
XX
SQ Sequence 553 BP; 159 A; 136 C; 128 G; 127 T; 3 other;

Query Match 19.2%; Score 492.4; DB 24; Length 553;
Best Local Similarity 97.5%; Pred. No. 2e-117;
Matches 541; Conservative 0; Mismatches 9; Indels 5; Gaps 4;

OY 1322 GGTGGTAAGCTAATGAGTCCGCGAAGTCGATGTCGTAATACTGACAGAGAA 1381
|||||
DB 1 GGTGCTAAGCTAATGAGTCCGCGAAGTCGATGTCGTAATACTGACAGAGAA 60
OY 1382 TAAATAATGATGACACGGGAGCAGTGGATCGCATGAAAAACAGTGTATCGTATCAA 1441
|||||
DB 61 TAAGAATGATGACACGGGAGCAGTGGATCGCATGAAAAACAGTGTATCGTATCAA 120
OY 1442 TATGGCCCACTCCACACAGAGAAATGATGACAGCCTCCGCACTCCGGAGCTGACGTG 1501
|||||
DB 121 TATGGCCCACTCCACACAGAGAAATGATGACAGCCTCCGCACTCCGGAGCTGACGTG 180
OY 1502 GGAGCGAGTACGTTCTCAGGTGACCATGTCATCTGGCCAGATGGCAACGAGTGTCTCT 1561
|||||
DB 181 GGAGCGAGTACGTTCTCAGGTGACCATGTCATCTGGCCAGAT-GNAAACGAGTGTCTCT 239

OY 1562 CCTGGCAGAGGCTGCTCTACTCAATTTGAGCTGCTCCACAGTTCACCACCTTGTCTGTG 1621
|||||
DB 240 CCTGGCAGAGGCTGCTCTACTCAATTTGAGCTGCTCCACAGTTCACCACCTTGTCTGTG 299
OY 1622 CATCACAGCCACACACAGAGCTTTGGCAGCTGATGAACTCTATATATGACCCGAGGGCG 1681
|||||
DB 300 CATCACAGCCACACACAGAGCTTTGGCAGCTGATGAACTCTATATATGACCCGAGGGCG 359
OY 1682 ATACAAGCAGAGTGTACTTGTCTCTCCTAAGAAA-TGGATGAATACGTGGCCAGCTTGC 1740
|||||
DB 360 ATACAAGCAGAGTGTACTTGTCTCTCCTAAGAAAATGGATGAATACGTGGCCAGCTTGC 419
OY 1741 ATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1800
|||||
DB 420 ATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 478
OY 1801 GACTCAACAAAAATGGGCCATTCAAACCTAATTATTACAGATATAATGACCATACTAC 1860
|||||
DB 479 GACTCAACAAAA--TGGGCCATTCAAACCTAATTATTACAGATATAATGACCATACTAC 536
OY 1861 CAAGGACCAGTCCAC 1875
|||||
DB 537 CAAGGACCAGTCCAC 551

RESULT 11
ABLO3692/c
ID ABL03692 standard; CDNA; 4406 BP.
XX
AC ABL03692;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 5558.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
PT P-PSDB; ABB59589.
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions
XX
PS Claim 1; SEQ ID NO 5558; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

QY 706 GTACACATCAGAGCCAGAGCGGTGTGATTGAGACACTCTGTGCCCCGCGGCTC 765
|||||
Db 502 GTACACATCAGAGCCAGAGCGGTGTGATTGAGACACTCTGTGCCCCGCGGCTC 443
QY 766 AGTGCCGCTGTCTGTCTGTAACTACTCACTCACTCAGATGAGTAGCTGCACACTGG 825
|||||
Db 442 AGTGCCGCTGTCTGTCTGTAACTACTCACTCACTCAGATGAGTAGCTGCACACTGG 383
QY 826 CTGAGGCTGAGTGTGAGTGTGCGTTGGAAGGCGAGTCAGAGATGACTTGTGTGT 885
|||||
Db 382 CTGAGGCTGAGTGTGAGTGTGCGTTGGAAGGCGAGTCAGAGATGACTTGTGTGT 323
QY 886 GTATTGACCGCTGTGTGACATGAGTGGTGCGACGCAACATGATCTGTGATGAGGG 945
|||||
Db 322 GTATTGACCGCTGTGTGACATGAGTGGTGCGACGCAACATGATCTGTGATGAGGG 263
QY 946 GAGACTTAACCCACTGGTATTATAAGATATCCAAACGTGTTAAGAGATCCGAGCA 1005
|||||
Db 262 GAGACTTAACCCACTGGTATTATAAGATATCCAAACGTGTTAAGAGATCCGAGCA 203
QY 1006 TTGTGGAAGAGAGCGTGTGCTGTCAAGGCTGTATCAGCTCTCCAAAGCTGGGAGC 1065
|||||
Db 202 TTGTGGAAGAGAGCGTGTGCTGTCAAGGCTGTATCAGCTCTCCAAAGCTGGGAGC 143
QY 1066 TCTGTGTTCGCGCATGAGCTCAATGATCTGTACCAACAGAGTTGATTAATCTGT 1125
|||||
Db 142 TCTGTGTTCGCGCATGAGCTCAATGATCTGTACCAACAGAGTTGATTAATCTGT 83
QY 1126 AC 1127
||
Db 82 AC 81

RESULT 13
ABL29786
ID ABL29786 standard; DNA; 5537 BP.
XX ABL29786;
AC
XX 26-MAR-2002 (first entry)
DT
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 40831.
DE
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
PT
XX
PS Claim 1; SEQ ID NO 40831; 21bp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (AB57737-AB572072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 5537 BP; 1632 A; 1054 C; 1158 G; 1693 T; 0 other;
Query Match 16.4%; Score 421.6; DB 23; Length 5537;
Best Local Similarity 59.78; Pred. No. 1.6e-98;
Matches 852; Conservative 0; Mismatches 439; Indels 137; Gaps 3;
QY 438 TCTTGTCTGCTCGATCTCAGAGTCTCCACTGACAGCTACAGTTCAGCTGCATCTAC 497
|||||
Db 2826 TCTCTCTGCGATTGAACCTCACTTAATGCTCACTGATCTCTTGTAGCCTGTTTACC 2885
QY 498 ACAGATAGCTCTGATGATGAGGTTTCTCCCGAGAGACAGCAACCAACTCCAGAGGC 557
|||||
Db 2886 GGAAGCTCCGACGAGAGAACGCTCTCCCGAGAGACCAATCAGCGCACTCCGCGCGA 2945
QY 558 AGCAGCAATTTCTGTGTGAAGAACATCAAGCAGGAGGAGTTGAGCGCGGAGATTGAG 617
|||||
Db 2946 GGCAGTACTTTTGGGTGAAGAGCATCTCGAAGAGTGTGAAAGCGGAGATCGAG 3005
QY 618 ATTGAGAGCAAGACATGTCTGCTGTATTTCACTCAGCAAGCTGCTCAGGGGAGAG 677
|||||
Db 3006 ATCGCCGAGTCGAGATGCGGGCATCATGACTCTGAGGAAAGAGCGAAGATGAGAG 3065
QY 678 CCCTTGCTGCTGCTAAATAGTGGGCTGTACACACATCAGCCAGACAGCGGTGTG 737
|||||
Db 3066 CCCCTAAGAGGTGCCAATATCTCGGATGACACCCAGTCAATGCTCAGTCGGCAGTCTG 3125
QY 738 ATTGAGACACTCTGTGCTGCTGGGGCTCAGTGGCGCTGCTGTCTGTATCATCTACTCA 797
|||||
Db 3126 ATGAGACCTCTGCTCAACTGCGGCGCCACAGTTGCTGCGCTGCTGCAACATTTATTC 3185
QY 798 ACTCAGATGAAGTACTGTCAGCACTGGCTGAGGCTGAGTGCAGTGTGCTTGAAG 857
|||||
Db 3186 ACACAAAGCAGATTGCCCGCTCTGCGAGAGCGGGAATTCGATCTTCGCTGGCGC 3245
QY 858 GCGAGTCAGAGATGACTTCTGTGTGTGATGACCGCTGTGTGACATGATGAGGTG 917
|||||
Db 3246 GGAGAGACGAGAGAGTCTGTGTGTGCTGTGACAGAGGCCATCTACTCCGACGCTG 3305
QY 918 CAGGCCACATGATCTCTGATGATGGGAGAGACTTAACCACTGGGTTTATAAGAGTAT 977
|||||
Db 3306 CAGCCGACCTTAATCTGAGCAGCGGCGGATGCCACGACCTTATGCTCAGAGAGTAC 3365
QY 978 CCAACGCTTTAAGAGATCCGAGGCTGTGGAAGAGAGCGTGAAGTGTGTACAGG 1037
|||||
Db 3366 CCGACTACTTCAAGGCCATTCGGGGCATCTGAGAGAAAGTGTGACCGGGGTGACCCG 3425
QY 1038 CTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTTCGGCCATGAAGCTCAATGATCT 1097
|||||
Db 3426 CTGTACATGCTGTCAAGGCGGAAACTTACTGTTCGGCCATCAACGTTAAGCACTCA 3485
QY 1098 GTTACCAACAGAGTTGATTAACCTGTGACTGTGCGGAGAAATCCATTTTGA----- 1150
|||||
Db 3486 GTGACCAAGAACAGTTGATTAACCTTCTACACAGTGTGCTGACTCCATCCTGACAGTGA 3545
QY 1151 ----- 1150
Db 3546 GCATAGATACATGATTAAGGAAGAGTACTATCTGTGCTGCTTCTAATAATTC 3605
QY 1151 -----TGCCCTGAAGAGACACAGATGTGATGTTGTGGGGAACAAGTGTGTGTG 1206
|||||
Db 3606 CTTCAGTCTGAACGACACCGATATATGTTTGGCGAAGCAAGGTGTGATCTGTG 3665
QY 1207 GCTATGTGAGTGAAGCAAGGCTGTGTGCTGCTCTTCAAGCTCTTGAGCAATGTCT 1266
|||||
Db 3666 GGTACGCTGATGTGGAAAGGGCTGTGCCAGTCTCCCTGAAGGGCCAGAGATGATGT 3725

Accession	Source	Sequence	Length
QY 1267	ACATTTACCGAAATCGACCCCATCTGTGCTCTGACGGCCCTGCATGATGGGTTGACGGGTGG	1326	
Db 3726	ATGTTACGGAGATGGATCCCATATGTGCTCTACAAAGCTGCCATGGATGGATTCGGGTGG	3785	
QY 1327	TAAAGCTAAATGAAGTCATCCGGCAAGTCGATGTCGTAATTAACCTTGCACAGAAATAAGA	1386	
Db 3786	TACGGCTCAACGAGGTTCATCAGGACGGTGGATGTGGTGTACGGCAACTGGAAACAAAA	3845	
QY 1387	ATGTAGTGCACGCGGAGCACCCTTGGATCGCATGAAAAACAGTTGTATCGTATGCATATGG	1446	
Db 3846	ATGTTATTTACCAAGGATCATCATGAATCGCATGAAGAATGGTTGTATCCTCTGCAATATGG	3905	
QY 1447	GCCACTCCACACAGAAATCGATGT-----1471	1471	
Db 3906	GACATTCCTGCTCGGAGATGTGATGTGATGATTCGGGGTGATTCATTTGTGGCCTTCA	3965	
QY 1472	-----GACCAGCCTCCGCACTCCGGAAGTACGTGGG	1503	
Db 3966	TATATCAAAAGGTTTTTAACCCACTTTAGATGGCTTGCAATACCCCGAGCTAACGTGGG	4025	
QY 1504	AGCGATACGTTCTCAGGTGGACCATGTCATCTGGCCAGATGGCAACGAGTTGTCTCC	1563	
Db 4026	AGCGTGTCCGTTCTCAAGTGATCATCATCAGTGGCCGAGCGCAGATGATCATTTTGC	4085	
QY 1564	TGGCAGAGGGTGTCTACTCAATTTGAGCTGTCCACAGTTCCACCTTGTCTGTCCA	1623	
Db 4086	TGCGCGAGGGAAGACTGGTGAATTTGTCTCTGTCACCATTTTGTCTTGTCTATCCG	4145	
QY 1624	TCACAGCCACACACAGCCTTTGGCAGCTGATAGAATCTATAATGCACCCGAGGGCGAT	1683	
Db 4146	TGGCTCATCCACCCAGGCTTTGGCCCTGATTTGAACCTCTCTCAGCGCC--AGGAAGAT	4202	
QY 1684	ACAAGCAGGATGTGTACTGCTCTCTTAAGAAATGATGAATACGTTG 1731		
Db 4203	ATAAGTCGATGTCTACTGCTGCTGCAAGAAATGGGTAGTTCTTAG 4250		
RESULT 14			
ID	AAZ80766/c		
XX	AAZ80766 standard; cDNA; 636 BP.		
XX	AAZ80766;		
DT	07-APR-2000 (first entry)		
XX			
DE	Human colon cancer cell line SW480 cDNA clone SEQ ID NO:850.		
XX			
KW	Human; gene expression product; diagnosis; tumour; colon cancer;		
KW	colorectal adenocarcinoma; cell line SW480; cell proliferation;		
KW	cytostatic; sarcoma; breast cancer; neoplasia; dysplasia;		
KW	hyperplasia; ds.		
XX			
OS	Homo sapiens.		
XX			
PN	W09964576-A2.		
XX			
PD	16-DEC-1999.		
XX			
PF	09-JUN-1999; 99WO-IB01062.		
XX			
PR	10-JUN-1998; 98US-0088801.		
XX			
PA	(FARB) BAYER CORP.		
XX			
PI	Endege WO, Steinmann KE, Astle JH, Burgess CC, Bushnell SE;		
PI	Carroll E, Catino TJ, Derti A, Ford DM, Lewis ME, Monahan JE;		
PI	Schlegel R;		
XX			
DR	WPI; 2000-087220/07.		
XX			
PT	Novel nucleic acids, used to develop products for the diagnosis and treatment of disorders involving unwanted cell proliferation,		

PT	particularly cancers, especially colon cancer
XX	
PS	Claim 15; Page 469; 469pp; English.
CC	AZ79917 to AZ80766 represent double stranded cDNA clones isolated from
CC	the human colorectal adenocarcinoma (colon cancer) cell line SW480. The
CC	cDNA clones can be used to generate antisense oligonucleotides which
CC	can be used for antisense therapy. Methods and products from the present
CC	invention can be used for identifying and/or classifying cancerous cells
CC	present in a human tumour, particularly in solid tumours, e.g.
CC	carcinomas and sarcomas, e.g. breast or colon cancers. The cDNA clones
CC	can be used for developing agents for the diagnosis and treatment of
CC	disorders involving unwanted cell proliferation, such as neoplasia,
CC	dysplasia or hyperplasia.
XX	
SQ	Sequence 636 BP; 156 A; 166 C; 137 G; 147 T; 30 other;
Query Match	15.9%; Score 406.4; DB 21; Length 636;
Best Local Similarity	99.5%; Pred. No. 4.5e-95;
Matches 418; Conservative	0; Mismatches 1; Indels 1; Gaps 1;
QY	707 TACACATCAGAGCCAGACAGC-GGTGTTGATTGAGACTCTGGCCCTGGGGCTC 765
Dd	420 TACACATCAGAGCCAGACAGCGGGTGTGATTGAGACACTCTGTCTGGGGCTC 361
QY	766 AGTCGGCTGCTGCTGTGAACAATACTCAACTCAGAATGAAGTAGCTGCAGCACTGG 825
Dd	360 AGTCGGCTGCTGCTGTGAACAATACTCAACTCAGAATGAAGTAGCTGCAGCACTGG 301
QY	826 CTGAGGCTGGAGTTGCAGTGTCGCTTGAAGGGCGAGTCAGAGATGACTTCTGTGTGT 885
Dd	300 CTGAGGCTGGAGTTGCAGTGTCGCTTGAAGGGCGAGTCAGAGATGACTTCTGTGTGT 241
QY	886 GTATTGACCGCTGTGTGAACATGATGGGTGGCAGGCCAATCATGATCCTGATGATGGGG 945
Dd	240 GTATTGACCGCTGTGTGAACATGATGGGTGGCAGGCCAATCATGATCCTGATGATGGGG 181
QY	946 GAGACTTAACCCACTGGGTTTAAAGAGTATCCAACGTTTAAAGAGATCCGAGGCA 1005
Dd	180 GAGACTTAACCCACTGGGTTTAAAGAGTATCCAACGTTTAAAGAGATCCGAGGCA 121
QY	1006 TTGTGGAAGAGAGCGTGAAGTGTGTACAGGCTGTATCAGCTCTCCAAAAGCTGGGAAGC 1065
Dd	120 TTGTGGAAGAGAGCGTGAAGTGTGTACAGGCTGTATCAGCTCTCCAAAAGCTGGGAAGC 61
QY	1066 TCTGTCTCCGGGCATGAAGCTCATATCTCTGTACCACAAAGAGTTGATACTTGT 1125
Dd	60 TCTGTCTCCGGGCATGAAGCTCATATCTCTGTACCACAAAGAGTTGATACTTGT 1
RESULT 15	
ID	AAC77823 standard; cDNA; 2200 BP.
AC	AAC77823;
DT	08-FEB-2001 (first entry)
DE	Human cancer associated gene sequence SEQ ID NO:217.
XX	
KW	Human; cancer associated gene; cancer antigen; detection; cancer;
KW	diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
KW	antidiabetic; antisthmatic; antirheumatic; antiarthritic; antiviral;
KW	antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
KW	dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
KW	vasotropic; antipsoriatic; angiogenic; gene therapy; inflammation;
KW	immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW	allergic reaction; graft versus host disease; organ rejection;
KW	haemostatic; thrombolytic; cardiovascular disorder; infection;
KW	neurological disease; drug screening; ss.
OS	Homo sapiens.
XX	

PN WO20005350-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US05882.
XX
PR 12-MAR-1999; 99US-0124270.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR MPI: 2000-587533/55.
DR P-PSDB; AAB43614.
XX
PT Novel isolated nucleic acids comprising sequences encoding peptides
PT useful for treating or diagnosing e.g. cancer -
XX
PS Claim 1; Page 787-788; 2352pp; English.
XX
CC AAC77607 to AAC78448 encode the human cancer associated proteins given
CC in AAB4398 to AAB44239. The proteins can have activities based on the
CC tissues and cells the genes are expressed in. Example of activities
CC include: cytostatic; proliferative; vulnerary; immunomodulator;
CC antidiabetic; antisthmatic; antineumatic; antiarthritic;
CC antiinflammatory; neuroprotective; cardiant; thrombolytic; coagulant;
CC dermatological; vasotropic; antipsoriatic and antiangiogenic. The
CC polynucleotides and polypeptides can be used for preventing, treating or
CC ameliorating medical conditions and diagnosing pathological conditions.
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat immune disorders by activating
CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells, to treat disorders of haematopoietic cells, autoimmune
CC disorders, allergic reactions, graft versus host disease and organ
CC rejection, modulate haemostatic or thrombolytic activity, modulate
CC inflammation, cancers, cardiovascular disorders, neurological disease and
CC bacterial or viral infections. The peptides, nucleotides, antibodies,
CC agonists and antagonists may be also be used in drug screens. AAC78449 to
CC AAC78457 and AAB44240 represent sequences used in the exemplification of
CC the present invention.
XX
SQ Sequence 2200 BP; 546 A; 578 C; 607 G; 466 T; 3 other;
Query Match 15.8%; Score 404; DB 21; Length 2200;
Best Local Similarity 58.4%; Pred. NO. 3.7e-94;
Matches 745; Conservative 0; Mismatches 525; Indels 6; Gaps 2;
QY 580 ACATCAAGCAGGAGATTTGAGCGCCGGAGATTTGAGATTGACAGACAGACATGTCG 639
DB 108 ACATCGGCTGGCTGGCTGGGAGCGCAAGCCCTGGACATGCTGAGAACGAGATGCCG 167
QY 640 CTCGTATTCTACAGGAACGCTGCTCAGGGGAGAGACCCCTGGCTGCTAAATAG 699
DB 168 GCTGATGCGTATGCGGAGCGGTACTCGGCTCCAGCCACTGAAGGGCGCCGCTATCG 227
QY 700 TGGGCTGTACACATCAGACAGCCAGACAGCGGTGTGATTGAGACACTGTCGCTGG 759
DB 228 CTGGCTGCTGACATGACCGTGGAGAGCGGCTCTCATTTGAGACCCCTGTCACCCCTGG 287
QY 760 GGGCTCAGTGCCTGCTGCTGTTGTAACATCTACTCACTCAGATGAGTAGCTGACG 819
DB 288 GTGCTGAGTGCAGTGTCCAGCTGCAACATCTCTCCACCCAGACCATGCGCGGCTG 347
QY 820 CACTGCTGAGGCTGAGTGTGAGTGTGCTGTTGGAAGGCGGAGTCAAGAGATGACTTCT 879
DB 348 CCATGCGCAAGGCTGGCATTCGGGTGTATGCTGGAAGGGCGAAGGAGAGAGATACC 407
QY 880 GGTGTGTATGACCGCTGTGTGACATGATGGGTGGCAGGCCAATGATCTGATG 939
DB 408 TGTGTGCTATGAGCAGACCCGTGTACTTCAAGGAGCGGCCCTCAACATGATTTGAGCG 467
QY 940 ATGGGGGAGACTTAACCCACTGGGTTTATAAGATATCCAAACGTTTGAAGAGATCC 999

DB 468 ACGGGGCGACCTTCAACCACTCATCCACACCAAGTACCCGACGCTTCTGCCAGGATCC 527
QY 1000 GAGCATTTGGAAGAGAGCGTGAAGTGTGTTACAGAGCTGTATCAGCTCTCCAAAGCTG 1059
DB 528 GAGCATCTCTGAGAGAGACACGACTGGGTTCACAACTCTACAAAGATGAGCCAATG 587
QY 1060 GGAAGCTCTGTGTCCGCCCATGAACGTCATGATCTCTTACCAACAGAGATTGATA 1119
DB 588 GGATCCTCAAGGTGCTGCCATCATGATGATGATGATGATGATGATGATGATGATGATG 647
QY 1120 ACTTGTAGTGTGCTGCCAGAAATCCATTTGGATGGCTTGAAGAGACACAGATGTGATGT 1179
DB 648 ACCCTATGCTGCTGCCGAGAGTCCCTCATGATGATGATGATGATGATGATGATGATG 707
QY 1180 TTGGTGGAAACAGTG 1239
DB 708 TTGCCGGCAAGTATGCGGTG 767
QY 1240 CTCTCAAGCTCTTGGAGCAATTTCTTACATTCAGAAATCGAACCCCATCTGTGCTGTGC 1299
DB 768 CCTGCGGGTGTTCGAGGCCCGCTCATCATCACCAGATGTACCCCATCAACGACATGC 827
QY 1300 AGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1359
DB 828 AGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 887
QY 1360 TCGTATTAATGTCACAGAAATTAAGATGTAGTACACGGGACCACTTGATCGCATGA 1419
DB 888 TCTTGTACACACACAGGCTGTATGTACATCATCTTGGCGGCACTTGAAGCATGA 947
QY 1420 AAAACAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1479
DB 948 AGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1007
QY 1480 TCCGCACTCCGAGTG 1539
DB 1008 TCAACGAGAACGCGGTGAGAGAGTGAACATCAAGCCGAGTGAACCGTATCGGTGA 1067
QY 1540 CAGATGCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1599
DB 1068 AGATGCGCGCGCATCATCTGCTGCGCGAGGCTGCGGTGCAACCTGGGTGTGCA 1127
QY 1600 CAG--TTCCACCTTG 1656
DB 1128 TGGGCCACCCAGCTTG 1187
QY 1657 AACTCTATATGACCCGAGGCGGATACAAAGAGATGTGTGTGTGTGTGTGTGTGTGTG 1716
DB 1188 AGCTGTGACCCATCCAGACAAG--TACCCCGTGGGTGTGTGTGTGTGTGTGTGTGTG 1244
QY 1717 TGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1776
DB 1245 TGGATGAGCAGTG 1304
QY 1777 CAGATGACCAAGCAAAATATCTGGAGCTCANCAAAATGGGCAATTCAAACCTAATTAT 1836
DB 1305 CTGAGAGCAAGCCAGTACCTGGGATGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1364
QY 1837 ACAGATACTAATGAC 1852
DB 1365 ACCGCTACTGAGAGCC 1380

Search completed: April 20, 2003, 14:39:52
Job time : 578 secs

QY 1741 ATTCGCATCATTTGATGCCACCCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1800
Db 1741 ATTCGCATCATTTGATGCCACCCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1800
QY 1801 GACTCAACAAAATGGGCCATTCAAACCTTAATTATACAGATACTAATGGACCATACTAC 1860
Db 1801 GACTCAACAAAATGGGCCATTCAAACCTTAATTATACAGATACTAATGGACCATACTAC 1860
QY 1861 CAAGACACAGTCCACCTGAAC 1920
Db 1861 CAAGACACAGTCCACCTGAAC 1920
QY 1921 TTTCTTCTTACTCCTTCTCTCTGATTTTTCCTAATAATTTCATTTCTGTTTTCATC 1980
Db 1921 TTTCTTCTTACTCCTTCTCTCTGATTTTTCCTAATAATTTCATTTCTGTTTTCATC 1980
QY 1981 TCATTAATCCAAAGTCTGACAGACACACAGAACTGCTTCATGGCTCTTAGATGAATA 2040
Db 1981 TCATTAATCCAAAGTCTGACAGACACACAGAACTGCTTCATGGCTCTTAGATGAATA 2040
QY 2041 GAAGTTCAGGTCCTCTCACTCTAGTCACTAAAGAGATTTTACTCCCCAGCCAGAGAA 2100
Db 2041 GAAGTTCAGGTCCTCTCACTCTAGTCACTAAAGAGATTTTACTCCCCAGCCAGAGAA 2100
QY 2101 GGTGATCTCTCTTTTACCATTTCTGGGAGCTTTAGTCTTAATTAGTACCTTATTACA 2160
Db 2101 GGTGATCTCTCTTTTACCATTTCTGGGAGCTTTAGTCTTAATTAGTACCTTATTACA 2160
QY 2161 GGAATGCTAAGGTACCTCTCTGTGGAACAATCTGCAATGCTAATCGCCTTAAAGA 2220
Db 2161 GGAATGCTAAGGTACCTCTCTGTGGAACAATCTGCAATGCTAATCGCCTTAAAGA 2220
QY 2221 GCCCATTTCTTACCTGCTGGAATCAGTGTCTTTCACCTTCTCAGAGAAGCAGGATGCT 2280
Db 2221 GCCCATTTCTTACCTGCTGGAATCAGTGTCTTTCACCTTCTCAGAGAAGCAGGATGCT 2280
QY 2281 ACCTACCCGGCAGGTAGGTAGATGTGGGTGCTGATGTTAATTTCCCTTAGAAGTCCA 2340
Db 2281 ACCTACCCGGCAGGTAGGTAGATGTGGGTGCTGATGTTAATTTCCCTTAGAAGTCCA 2340
QY 2341 AGCCCTGTTCTGCTGCTAAGGTGATGTCCAGTTCAGAGATGCTAATATGAGCATGG 2400
Db 2341 AGCCCTGTTCTGCTGCTAAGGTGATGTCCAGTTCAGAGATGCTAATATGAGCATGG 2400
QY 2401 CTTGTTAAGATCAGAGGCCCACTTGATTTATAGATAGCCCTTCTCCACTCCACCA 2460
Db 2401 CTTGTTAAGATCAGAGGCCCACTTGATTTATAGATAGCCCTTCTCCACTCCACCA 2460
QY 2461 GACTGCTCATTTTTCAGAGTTTAACTAGACTACACTCTAATTGAGTTAATTGTCC 2520
Db 2461 GACTGCTCATTTTTCAGAGTTTAACTAGACTACACTCTAATTGAGTTAATTGTCC 2520
QY 2521 TCTAGGATTTATTTCTGTTGCCAAAAA 2563
Db 2521 TCTAGGATTTATTTCTGTTGCCAAAAA 2563

RESULT 2
AF315687 2677 bp mRNA linear PRI 22-MAR-2002
LOCUS AF315687
DEFINITION Homo sapiens S-adenosylhomocysteine hydrolase-like protein mRNA,
complete cds.
ACCESSION AF315687
VERSION AF315687.1 GI:16588686
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2677)
Dekker,J.W., Budhia,S., Angel,N.Z., Cooper,B.J., Clark,G.J.,
Hart,D.N. and Kato,M.

TITLE Identification of an S-adenosylhomocysteine hydrolase-like
JOURNAL transcript induced during dendritic cell differentiation
MEDLINE Immunogenetics 53 (12), 993-1001 (2002)
PUBMED 21901265
REFERENCE 2 (bases 1 to 2677)
AUTHORS Dekker,J.W., Budhia,S., Angel,N.Z., Cooper,B.J., Clark,G.J.,
Hart,D.N.J. and Kato,M.
TITLE Direct Submission
JOURNAL Submitted (23-Oct-2000) Dendritic Cell Research, Mater Medical
Research Institute, Level 3, Aubigny Place, South Brisbane,
Queensland 4101, Australia
FEATURES
source location/Qualifiers
1..2677
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p21"
/clone="211(1)B"
/cell_line="Hodgkin's disease cell line L428"
1..114
/note="derived from 5' RACE"
369..1961
/note="DCAL; dendritic cell expressed AHCY-like protein;
expressed in dendritic cells in blood, Langerhans cells
and dermal dendritic cells; mRNA increased during monocyte
differentiation to monocyte-derived dendritic cells in the
presence of GM-CSF and IL-4"
/codon_start=1
/product="S-adenosylhomocysteine hydrolase-like protein"
/protein_id="AA126869.1"
/db_xref="GI:16588687"
/translation="MSMPDAMPLPGVGEELKQAKEDAEKYSFMATVTKAPKQIOF
ADDMQEFTEPRTKGRSLRSISOSSTDSYSSASYSDDSDDEVSPREKQOTNSKS
SNFCVKNIKQAEGRREIEIAQDMASLISLRKAGEKPLAGAKIVGCTHTAQTAV
LIETICALGACQMSACNISTQNEVAALAEAGVAFAMKGESEDDFWCIDRCVNM
DGWQANMILDDGDLTHWVYKYPNVFKIRGIVESVTGVHRLYQLSKAGKLCVPM
NVNDSVTKQKEDNLCCRESILDLKRTDVMFGKQYVCGYGEVKGCCALRALG
AIVYITEIDPICALQACMDGERVVKLINEVIRVDVITCTGNKNVTRHLDRKNSC
IVCNMGSNTEIDVTSRLTPELTWERVSQVDHYWFDGKRVLLAEGRLNLSCTV
PTFVLSTATATQALALIELYNAPEGRYKQDVLLPKMDEVVASLHLSFDAHLTELT
DDQAKYLGLNKNQEPFKNYRYR"
BASE COUNT 657 a 648 c 708 g 664 t
ORIGIN
Query Match 100.0%; Score 2563; DB 9; Length 2677;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCCGGGCAGGTGAGAGCTGAGCTGCTCTTCTGTTCTGTTGAGCCACCGTGGCT 60
Db 115 GCCCGGGCAGGTGAGAGCTGAGCTGCTCTTCTGTTCTGTTGAGCCACCGTGGCT 174
QY 61 GTCCGGCTGCTTGGGCTGCCGACAGACAAGCGGTGGGCCACAGCACTCAGAAAGCCGA 120
Db 175 GTCCGGCTGCTTGGGCTGCCGACAGACAAGCGGTGGGCCACAGCACTCAGAAAGCCGA 234
QY 121 CGCAGCTCGACGAGGGGGCCGAGAGGGTGGCGATCGCGTGTGAGAGGGCGCCGGCG 180
Db 235 CGCAGCTCGACGAGGGGGCCGAGAGGGTGGCGATCGCGTGTGAGAGGGCGCCGGCG 294
QY 181 GGGCAGGGGGGGGGGGGGGGGAGAGAGAGGGGGGGGGGGGGGGGGGGTACGGCTGGCC 240
Db 295 GGGCAGGGGGGGGGGGGGGGGAGAGAGAGGGGGGGGGGGGGGGGGGGTACGGCTGGCC 354
QY 241 GGGCGGGGGGGGGGAGATGTCAGATGCTGACGCGATGCGGCTGCGGGGGTGGGAGGAGC 300
Db 355 GGGCGGGGGGGGGGAGATGTCAGATGCTGACGCGATGCGGCTGCGGGGGTGGGAGGAGC 414
QY 301 TGAAGCAGGCCAAGAGATGAGAGCGGAGAGAGTACTCTTCTCATGGCCACCGTCACCA 360
Db 415 TGAAGCAGGCCAAGAGATGAGAGCGGAGAGAGTACTCTTCTCATGGCCACCGTCACCA 474

QY 361 AGCGGCCCAAGAGCAATCCAGTTGCTGATGACATGCAGAGTTCCACCAATTCCCA 420
DB 475 AGCGGCCCAAGAGCAATCCAGTTGCTGATGACATGCAGAGTTCCACCAATTCCCA 534
QY 421 CCAAACTGGCCGAAGATCTTGTCTGCTGCATCTCACAGTCTCCACTGCAGACTACA 480
DB 535 CCAAACTGGCCGAAGATCTTGTCTGCTGCATCTCACAGTCTCCACTGCAGACTACA 594
QY 481 GTTCAGCTGCATCTCACAGATAGCTGTGATGATGAGTTTCTCCCGAGAGACAGC 540
DB 595 GTTCAGCTGCATCTCACAGATAGCTGTGATGATGAGTTTCTCCCGAGAGACAGC 654
QY 541 AAACCACTCCAAGGCGAGCAAAATTCTGTGGAAGACATCAAGCAGGAGAAATTG 600
DB 655 AAACCACTCCAAGGCGAGCAAAATTCTGTGGAAGACATCAAGCAGGAGAAATTG 714
QY 601 GACGCCGGAGATTGAGATTGCAGACAAGACATGTCTGCTGTGATTCTACAGAAAC 660
DB 715 GACGCCGGAGATTGAGATTGCAGACAAGACATGTCTGCTGTGATTCTACAGAAAC 774
QY 661 GTGCTCAGGGGAGAGGCCCTTGGCTGTGCTAAATAGTGGGCTGTACACATCACAG 720
DB 775 GTGCTCAGGGGAGAGAGGCCCTTGGCTGTGCTAAATAGTGGGCTGTACACATCACAG 834
QY 721 CCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGGGGGCTCAGTGGGCTGTCTG 780
DB 835 CCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGGGGGCTCAGTGGGCTGTCTG 894
QY 781 CTGTGACATCTACTCAACTCAGAAATGAGTAGCTGCAGCACTGGCTGAGGCTGAGTTG 840
DB 895 CTGTGACATCTACTCAACTCAGAAATGAGTAGCTGCAGCACTGGCTGAGGCTGAGTTG 954
QY 841 CAGTGTGCTTGAAGGGCGAGTCAGAAAGATGACTTCTGTGTGTTGACCGCTGTG 900
DB 955 CAGTGTGCTTGAAGGGCGAGTCAGAAAGATGACTTCTGTGTGTTGACCGCTGTG 1014
QY 901 TGAACATGATGGGTGGCAGGCCAATGATCTGATGATGGGGAGCTTAACCCACT 960
DB 1015 TGAACATGATGGGTGGCAGGCCAATGATCTGATGATGGGGAGCTTAACCCACT 1074
QY 961 GGGTTTATAGAAGTATCCAAAGCTGTTAAGAAGATCCGAGCATTTGGAAGAGCG 1020
DB 1075 GGGTTTATAGAAGTATCCAAAGCTGTTAAGAAGATCCGAGCATTTGGAAGAGCG 1134
QY 1021 TGACTGCTGTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTCCGGCCA 1080
DB 1135 TGACTGCTGTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTCCGGCCA 1194
QY 1081 TGAACGTCAATGATCTGTATACCAAAACAGAAATTGATACTTGTGCTCCGAGAA 1140
DB 1195 TGAACGTCAATGATCTGTATACCAAAACAGAAATTGATACTTGTGCTCCGAGAA 1254
QY 1141 CCATTTTGGATGGCCTGAAGAGGACACAGATGTGATGTTGGGAAACAAGTGTGG 1200
DB 1255 CCATTTTGGATGGCCTGAAGAGGACACAGATGTGATGTTGGGAAACAAGTGTGG 1314
QY 1201 TGTGTGCTATGTGAGTAGGCAAGGCTGTGCTGTCTCAAAAGCTTTGAGCAA 1260
DB 1315 TGTGTGCTATGTGAGTAGGCAAGGCTGTGCTGTCTCAAAAGCTTTGAGCAA 1374
QY 1261 TTGTCTACATTACCGAATTCGACCCCATCTGTGCTGTGAGGCTGCATGGATGGTTCA 1320
DB 1375 TTGTCTACATTACCGAATTCGACCCCATCTGTGCTGTGAGGCTGCATGGATGGTTCA 1434
QY 1321 GGGTGTAAAGCTAAATGAAGTCAATCCGGCAAGTGTGTAATACTTGACAGGAA 1380
DB 1435 GGGTGTAAAGCTAAATGAAGTCAATCCGGCAAGTGTGTAATACTTGACAGGAA 1494
QY 1381 ATAAGATGTAGTAGACAGGAGCACTTGATGCATGAATAACAGTTGTATCGTATGCA 1440
DB 1495 ATAAGATGTAGTAGACAGGAGCACTTGATGCATGAATAACAGTTGTATCGTATGCA 1554
QY 1441 ATATGGCCACTCCAACAGAAATCGATGTGACCAGCCTCCGCACTCCGAGCTGACGT 1500

DB 1555 ATATGGCCACTCCAACACAGAAATCGATGTGACCAGCCTCCGCACTCCGAGCTGACGT 1614
QY 1501 GGGAGCGAGTACGTTCTCAGGTGACCATGTCAATCTGGCCAGATGGCAACGAGTTGTCC 1560
DB 1615 GGGAGCGAGTACGTTCTCAGGTGACCATGTCAATCTGGCCAGATGGCAACGAGTTGTCC 1674
QY 1561 TCCTGGCAGAGGGTGTCTACTCAATTTGAGCTGTCCACAGTTCACACCTTTGTCTGT 1620
DB 1675 TCCTGGCAGAGGGTGTCTACTCAATTTGAGCTGTCCACAGTTCACACCTTTGTCTGT 1734
QY 1621 CCATCACAGCCACAACACAGGCTTTGGCAGCTGATGAACCTATATATGACCCGAGGGGC 1680
DB 1735 CCATCACAGCCACAACACAGGCTTTGGCAGCTGATGAACCTATATATGACCCGAGGGGC 1794
QY 1681 GATACAGCAGATGTGTACTTGTCTTCTTAAGAAATGATGAATACGTTGCCAGCTTGC 1740
DB 1795 GATACAGCAGATGTGTACTTGTCTTCTTAAGAAATGATGAATACGTTGCCAGCTTGC 1854
QY 1741 ATCTGCCATCATTTGATGCCCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1800
DB 1855 ATCTGCCATCATTTGATGCCCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1914
QY 1801 GACTCAACAAAATGGGCCATTCBAACCTAATTTACAGATACTAATGACCATACTAC 1860
DB 1915 GACTCAACAAAATGGGCCATTCBAACCTAATTTACAGATACTAATGACCATACTAC 1974
QY 1861 CAAGGACAGTCCACCTGAACACACACACTCTAAGAAATATTTTAAAGATACCTTTAT 1920
DB 1975 CAAGGACAGTCCACCTGAACACACACACTCTAAGAAATATTTTAAAGATACCTTTAT 2034
QY 1921 TTTCTTCTTACTCTTCTCTCTGATTTTCTCTAATTTTCAATCTTGTGTTTCTATC 1980
DB 2035 TTTCTTCTTACTCTTCTCTCTGATTTTCTCTAATTTTCAATCTTGTGTTTCTATC 2094
QY 1981 TCATATCCAAAGTCTGACAGACACACAGAACTGCTTCATGCTCTTAGATGAATA 2040
DB 2095 TCATATCCAAAGTCTGACAGACACACAGAACTGCTTCATGCTCTTAGATGAATA 2154
QY 2041 GAAGTTCAGGGTCCCTCACTCTAGTCACTAAGAAAGATTTTACTCCCCAGCCAGAAA 2100
DB 2155 GAAGTTCAGGGTCCCTCACTCTAGTCACTAAGAAAGATTTTACTCCCCAGCCAGAAA 2214
QY 2101 GGTGATCTTCTCTTTACCATTCTGGGAGCTTTAGTCTTAATTAGTACCTTAATAACA 2160
DB 2215 GGTGATCTTCTCTTTACCATTCTGGGAGCTTTAGTCTTAATTAGTACCTTAATAACA 2274
QY 2161 GGAATGCTAAGGTACCTTCTGTGGAACAACTGCAATGTCTAATCGCCTTAAGAA 2220
DB 2275 GGAATGCTAAGGTACCTTCTGTGGAACAACTGCAATGTCTAATCGCCTTAAGAA 2334
QY 2221 GCCCATTTCTTAGCTGTGAATCAGTGTCTTCACTCTTCAAGAGAGCAGGATGCT 2280
DB 2335 GCCCATTTCTTAGCTGTGAATCAGTGTCTTCACTCTTCAAGAGAGCAGGATGCT 2394
QY 2281 ACCTACCCGGCAGGTAGGTAGATGTGGGTGTCATGTTAATTTCCCTTAGAAGTTCCA 2340
DB 2395 ACCTACCCGGCAGGTAGGTAGATGTGGGTGTCATGTTAATTTCCCTTAGAAGTTCCA 2454
QY 2341 AGCCCTGTTTCTGCTGAAGGTGTATGCCAGTTCAGAGATGTGTATATGACATGG 2400
DB 2455 AGCCCTGTTTCTGCTGAAGGTGTATGCCAGTTCAGAGATGTGTATATGACATGG 2514
QY 2401 CTGTTAAGATCAGGAGGCCCATTTGATTAATAGCCCTTCTCCACTCCACCA 2460
DB 2515 CTGTTAAGATCAGGAGGCCCATTTGATTAATAGCCCTTCTCCACTCCACCA 2574
QY 2461 GACTTGTCAATTTTTCGAGTTTAACTAGACTACACTATTTGAGTTAATTTGTCC 2520
DB 2575 GACTTGTCAATTTTTCGAGTTTAACTAGACTACACTATTTGAGTTAATTTGTCC 2634
QY 2521 TCTAGATTTATTTCTGTGTCCAAAAAATAAAAAA 2563


```

Db 1081 CAATGATTCCTGTACCAACAGAGATTGATACTGTACTGCTGCCGAGATCCATT 1140
QY 1148 GGATGGCCTGAGAGAGACCACAGATGTGATGTTGGTGGAAACAAGTGTGTGTGG 1207
Db 1141 GGATGGCCTGAGAGAGACCACAGATGTGATGTTGGTGGAAACAAGTGTGTGTGG 1200
QY 1208 CTATGGTGAAGTAGGCAAGGGCTGCTGTGCTGCTCAAAAGCTCTTGAGCAATTGTCTA 1267
Db 1201 CTATGGTGAAGTAGGCAAGGGCTGCTGTGCTGCTCAAAAGCTCTTGAGCAATTGTCTA 1260
QY 1268 CATTACCGAAATCGACCCCATCTGTGCTGTCGAGGCTGCATGATGGTTCAGGGTGT 1327
Db 1261 CATTACCGAAATCGACCCCATCTGTGCTGTCGAGGCTGCATGATGGTTCAGGGTGT 1320
QY 1328 AAAGCTAAATGAAGTCATCCGGCAAGTCGATGTCGTAATACTTGACACAGAAATAGAA 1387
Db 1321 AAAGCTAAATGAAGTCATCCGGCAAGTCGATGTCGTAATACTTGACACAGAAATAGAA 1380
QY 1388 TGTAGTGACAGGAGACACTTGGATCGCATGAAAAACAGTTGTATCGTATGCAATATGG 1447
Db 1381 TGTAGTGACAGGAGACACTTGGATCGCATGAAAAACAGTTGTATCGTATGCAATATGG 1440
QY 1448 CCACCTCCACACAGAAATCGATGTGACACAGCCTCCGCACTCCGGAGCTGACGTGGAGCG 1507
Db 1441 CCACCTCCACACAGAAATCGATGTGACACAGCCTCCGCACTCCGGAGCTGACGTGGAGCG 1500
QY 1508 AGTACGTTCTCAGGTGGACCATGTCTATCTGGCCAGATGGCAAAAGAGTTGTCTCTGCGC 1567
Db 1501 AGTACGTTCTCAGGTGGACCATGTCTATCTGGCCAGATGGCAAAAGAGTTGTCTCTGCGC 1560
QY 1568 AGAGGTCGCTACTCAATTTGAGCTGCTCCACAGTCCCACTTTGTTCTGTCCATCAC 1627
Db 1561 AGAGGTCGCTACTCAATTTGAGCTGCTCCACAGTCCCACTTTGTTCTGTCCATCAC 1620
QY 1628 AGCCACAACACAGGCTTTGGCACTGATAGAACTCTAATAATGCAACCCGAGGGCGATACAA 1687
Db 1621 AGCCACAACACAGGCTTTGGCACTGATAGAACTCTAATAATGCAACCCGAGGGCGATACAA 1680
QY 1688 GCAGGATGTGACTTCTCTCTTAAGAAATGATGAATACGTTGCCAGCTTGCACTGCC 1747
Db 1681 GCAGGATGTGACTTCTCTCTTAAGAAATGATGAATACGTTGCCAGCTTGCACTGCC 1740
QY 1748 ATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAA 1807
Db 1741 ATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAA 1800
QY 1808 CAAAAATGGCCATTCAAACCTAATTATTACAGACTAATGACCATACTACCAAGGAC 1867
Db 1801 CAAAAATGGCCATTCAAACCTAATTATTACAGACTAATGACCATACTACCAAGGAC 1860
QY 1868 CAGTCCACCTGAACACACACACTCTAAGAAATATTTTAAGATAACTTTATTTCTTC 1927
Db 1861 CAGTCCACCTGAACACACACACTCTAAGAAATATTTTAAGATAACTTTATTTCTTC 1920
QY 1928 TTACTCCTTCTCTCTGATTTTTCCTATATTCATTCATCTGTTTTCATCTCATTA 1987
Db 1921 TTACTCCTTCTCTCTGATTTTTCCTATATTCATTCATCTGTTTTCATCTCATTA 1980
QY 1988 CCAAGTTCTGAGACACACAGAACTGCTTCATGCTCTTAGATGAATAGAAAGTTC 2047
Db 1981 CCAAGTTCTGAGACACACAGAACTGCTTCATGCTCTTAGATGAATAGAAAGTTC 2040
QY 2048 AGGTCCTCTCACTAGTCACTAAGAAAGATTTTCTCCAGCCAGAAAGGTGAT 2107
Db 2041 AGGTCCTCTCACTAGTCACTAAGAAAGATTTTCTCCAGCCAGAAAGGTGAT 2100
QY 2108 CTCTCTTACCAATTTCTGGGACTTAGTCTTAATTAGGTACCTTATTAACAGAAATG 2167
Db 2101 CTCTCTTACCAATTTCTGGGACTTAGTCTTAATTAGGTACCTTATTAACAGAAATG 2159
QY 2168 CTAAGGTACCTTCTCTGTGGAACAATCTGCAATGTCTAAATCGCCCTAAAGAGCCCAT 2227
Db 2160 CTAAGGTACCTTCTCTGTGGAACAATCTGCAATGTCTAAATCGCCCTAAAGAGCCCAT 2219

```

```

QY 2228 TCTTAGCTGTGAATCAGTCTCTTTCACTTCTTCAGAGACAGGGATGTACTACC 2287
Db 2220 TCTTAGCTGTGAATCAGTCTCTTTCACTTCTTCAGAGACAGGGATGTACTACC 2279
QY 2288 CGGACAGTAGTGTAGTGGGTGGTGCATGTTAATTTCCCTTAGAAGTTCGAAGCCCTG 2347
Db 2280 CGGACAGTAGTGTAGTGGGTGGTGCATGTTAATTTCCCTTAGAAGTTCGAAGCCCTG 2339
QY 2348 TTCTGCGTAAAGTGTGTATGTCCAGTTCAGAGATGTGTATATGACATGGCTGTTA 2407
Db 2340 TTCTGCGTAAAGTGTGTATGTCCAGTTCAGAGATGTGTATATGACATGGCTGTTA 2399
QY 2408 AGATCAGAGAGCCCACTTGATTTATAGTATAGCCCTTCCCTCAGTCCACAGACTTGC 2467
Db 2400 AGATCAGAGAGCCCACTTGATTTATAGTATAGCCCTTCCCTCAGTCCACAGACTTGC 2459
QY 2468 TCATTTTTCGAGTTTAACTAGACTACACTCTATT 2503
Db 2460 TCATTTTTCGAGTTTAACTAGACTACACTCTATT 2495

```

```

RESULT 4
LOCUS BC007576 2526 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, S-adenosylhomocysteine hydrolase-like 1, clone
ACCESSION MG:15558 IMAGE:3139729, mRNA, complete cds.
VERSION BC007576 GI:14043176
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.

```

```

REFERENCE 1 (bases 1 to 2526)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (10-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

```

```

REMARK NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca

```

```

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susana Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline
Scheln, Duane Smalus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

```

```

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 22 Row: h Column: 21
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, similarity but not identity to protein.

```

```

FEATURES
location/Qualifiers
1..2526

```

```

/organism="Homo sapiens"
/db_xref="LocustID:10768"
/db_xref="taxon:9606"
/clone="MGC:15558 IMAGE:3139729"
/tissue.type="Placenta, choriocarcinoma"
/clone_lib="NIH_MGC_21"

```


QY	1913	ACTTTTATTTCTCTTCTTACTCCCTTTCCCTCTGTGATTTTCTTCTATAATTTCATTCTTGTT	1972
Db	1876	ACTTTTATTTCTTCTTCTTACTCCCTTTCCCTCTGTGATTTTCTTCTATAATTTCATTCTTGTT	1935
QY	1973	TTTTCATCTCATTTATCCAGTTCCTGACAGACCACACAGAACTTGCTTCATGGCTCTTAG	2032
Db	1936	TTTTCATCTCATTTATCCAGTTCCTGACAGACCACACAGAACTTGCTTCATGGCTCTTAG	1995
QY	2033	ATGAATAGAGTTCAGGGTCCCTCACTCTAGTCACCTAAAGAAGATTTTACTCCCCAG	2092
Db	1996	ATGAATAGAGTTCAGGGTCCCTCACTCTAGTCACCTAAAGAAGATTTTACTCCCCAG	2055
QY	2093	CCCAGAAAGGTGATTTCTCTCTTTACCATTTCTGGGACCTTAGTCTTAATTAGGTACCT	2152
Db	2056	CCCAGAAAGGTGATTTCTCTCTTTACCATTTCTGGGACCTTAGTCTTAATTAGGTACCT	2114
QY	2153	TATTAACAGGAATGCTAAGGTACCTTCTCTGTGGAACAATCTGCAATGTCTAAATCGCC	2212
Db	2115	TATTAACAGGAATGCTAAGGTACCTTCTCTGTGGAACAATCTGCAATGTCTAAATCGCC	2174
QY	2213	TTAAAGAGCCCATTTCTTAGCTGCTGAATCAGTGCCTTTCACTTCTTCAGAGAAGCA	2272
Db	2175	TTAAAGAGCCCATTTCTTAGCTGCTGAATCAGTGCCTTTCACTTCTTCAGAGAAGCA	2234
QY	2273	GGGATGCTACCTACCCGGCAGGTAGGTTAGATGTGGGTGGTGCATGTTAATTTCCCTTAG	2332
Db	2235	GGGATGCTACCTACCCGGCAGGTAGGTTAGATGTGGGTGGTGCATGTTAATTTCCCTTAG	2294
QY	2333	AAGTTCGAAGCCCTGTTCCTGCTGAAGGTGTTATGTCAGTTCAGAGATGTGTATAAT	2392
Db	2295	AAGTTCGAAGCCCTGTTCCTGCTGAAGGTGTTATGTCAGTTCAGAGATGTGTATAAT	2354
QY	2393	GAGCATGGCTGTGTAAGATCAGAGAGCCCACTTGATTTATAGTATAGCCCTTCCCTCCAC	2452
Db	2355	GAGCATGGCTGTGTAAGATCAGAGAGCCCACTTGATTTATAGTATAGCCCTTCCCTCCAC	2414
QY	2453	TCCCACACAGACTTGCTCATTTTTCGAGTTTAACTAGACTACACTCTATT	2503
Db	2415	TCCCACACAGACTTGCTCATTTTTCGAGTTTAACTAGACTACACTCTATT	2465
RESULT 5			
LOCUS	BC010681	2583 bp mRNA linear	PRI 12-JUL-2001
DEFINITION		Homo sapiens, S-adenosylhomocysteine hydrolase-like 1, clone	
ACCESSION	BC010681		
VERSION	BC010681.1	GI:14715037	
KEYWORDS		MGC.	
SOURCE		Homo sapiens.	
ORGANISM		Homo sapiens.	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	
TITLE		1 (bases 1 to 2583)	
JOURNAL		Strausberg, R.	
		Submitted (10-JUL-2001) National Institutes of Health, Mammalian	
		Gene Collection (MGC), Cancer Genomics Office, National Cancer	
		Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,	
		USA	
REMARK		NIH-MGC Project URL: http://mgc.ncl.nih.gov	
COMMENT		Contact: MGC help desk	
		Email: cgaps-remail.nih.gov	
		Tissue Procurement: ATCC	
		CDNA Library Preparation: Life Technologies, Inc.	
		CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)	
		DNA Sequencing by: Baylor College of Medicine Human Genome	
		Sequencing Center	
		Center code: BCM-HGSC	
		Web site: http://www.hgsc.bcm.tmc.edu/cdna/	
		Contact: villalob@bcm.tmc.edu	
		Villalob, D.R., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,	

A.M., Holloway, M., Telford, B, Hodgson, A., Bouck, J., Yu, W.,
Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>
Series: IRAK Plate: 13 Row: b Column: 21
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 5729723.

FEATURES	Location/Qualifiers
source	1. .2583

```

/organism="Homo sapiens"
/db_xref="locusid:10768"
/db_xref="taxon:9606"
/clone="MGC:8936 IMAGE:3853747"
/tissue_type="Colon, adenocarcinoma"
/clone_1b="NH_MGC_65"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
367. .1869

```

```

/product="S-adenosylhomocysteine hydrolase-like 1"
/protein_id="AAH10681.1"
/db_xref="GI:14715038"
/translation="MATVTAKAPKKQIQPADDMQEFKFPPTKTRGRSLRSISQSSSTDS
YSSASAYTSDSDDEVSPREKQOTNSGSSNECVKNIKQAEFGRAIEIEAQDMSALIS
LKRRAQGEKPLAGAKIVGCTHITAQTAVLIELTLCALGAQCRRWSACNISTONEVAAL
AEAGVAVFAWKGESEDDFWWCIDRCVNMNDGQAMNILLDDGGDLTHVYKPPNVFKKI
RGIVEESVTGVHRLYOLSKAGKLCVPMNVNDSVTKQKFDNLYCCRESLIDGLKRTTD
VNEGKQVVCVGYGEVGGKCCALKALGAIVITEIDPICALQACDGEFRVYKLENI
RQVDVITICTGNKNVTRREHLDRKNSCIVCNMGHSNTEIDVTSLRTPELTWERVRSQ
VDHVTMPDGKRVLLAEGRLNLTSCSPTEFLSITATQALALIELYNAPEGRYKQD
VILLPKMDEYVASLHLPSEDAHLTELTDQAKYGLGNKNGPFKPNRYR"

```

Query Match	83.0%;	Score 2127;	DB 9;	Length 2583;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 2497;	Conservative	0;	Mismatches 5;	Indels 1;
				Gaps 1.

QY	1	GGGGCGGGCAGGTCGGAGCTCGGAGCTGCTGCTTCTGTGTTCTCTGTGTGGCCACCCTGCT	60
Db	23	GGGGCGGGCAGGTCGGAGCTCGGAGCTGCTGCTTCTGTGTTCTCTGTGTGGCCACCCTGCT	82
QY	61	GTCCGGCTGCTTGGGCTGCCCCAACAAGACGAGCGGTGGGCCACAGCACCTCAGAAAGCCGA	120
Db	83	GTCCGGCTGCTTGGGCTGCCCCAACAAGACGAGCGGTGGGCCACAGCACCTCAGAAAGCCGA	142
QY	121	CGCAGCTCGACGCGAGGGGGCCGACAGAGGGTGGGGCATCGCTGTGCGAGAGGGCCCGCCG	180
Db	143	CGCAGCTCGACGCGAGGGGGCCGACAGAGGGTGGGGCATCGCTGTGCGAGAGGGCCCGCCG	202
QY	181	GGCGAGGGCGGGCGGGCGCCAGAGGGGGAAGAGCGGGGGCGGGGTACGCCCTGGCC	240
Db	203	GGCGAGGGCGGGCGGGCGCCAGAGGGGGAAGAGCGGGGGCGGGGTACGCCCTGGCC	262
QY	241	GGGGCGGGCGGGGAATGTCGATGCTTGACGCGGATGCCGTGCCCGGGGGTGGGGAGGAGC	300
Db	263	GGGGCGGGCGGGGAATGTCGATGCTTGACGCGGATGCCGTGCCCGGGGGTGGGGAGGAGC	322
QY	301	TGAAGCAGGCCAAGGAGATCGAGGACGCCGAGAAAGTACTCCTTCATGGCCACCGTCACCA	360
Db	323	TGAAGCAGGCCAAGGAGATCGAGGACGCCGAGAAAGTACTCCTTCATGGCCACCGTCACCA	382
QY	361	AGGCGCCCCAAGAAGCAATTCAGTTGCTGATGACATGCAGAGTTACACAAATTCGCCA	420
Db	383	AGGCGCCCCAAGAAGCAATTCAGTTGCTGATGACATGCAGAGTTACACAAATTCGCCA	442
QY	421	CCAAAACTGGCCGAAGATCTTGTCTGCTCGATCTCACAGTCTCTCCACTGACAGCTACA	480
Db	443	CCAAAACTGGCCGAAGATCTTGTCTGCTCGATCTCACAGTCTCTCCACTGACAGCTACA	502
QY	481	GTTCAGCTGCATCCTACACAGATAGCTCTGATGATGAGGTTTCTCCCGAGAGAAGCAGC	540

Db 503 GTTCAGCTGCATCCTACACAGATAGCTGTGATGAGGTTTCTCCCGAGAGAAGCAGC 562
QY 541 AAACCACTCCAAGGGCAGCAGCAATTTCTGTGTGAAGACATCAAGCAGAGAAATTG 600
Db 563 AAACCACTCCAAGGGCAGCAGCAATTTCTGTGTGAAGACATCAAGCAGAGAAATTG 622
QY 601 GACGCCGGAGATTGAGATTGCAGAGCAAGACATGTCTCTGTGATTCTACTCAGAAAC 660
Db 623 GACGCCGGAGATTGAGATTGCAGAGCAAGACATGTCTCTGTGATTCTACTCAGAAAC 682
QY 661 GTGCTCAGGGGGAGAAAGCCCTGTGGCTGTCTAAATAGTGGGCTGTACACATCAGAG 720
Db 683 GTGCTCAGGGGGAGAAAGCCCTGTGGCTGTCTAAATAGTGGGCTGTACACATCAGAG 742
QY 721 CCCAGACAGCGGTGTGATTGAGACACTCTGTGCTGGGCTGAGTGGCGCTGTCTG 780
Db 743 CCCAGACAGCGGTGTGATTGAGACACTCTGTGCTGGGCTGAGTGGCGCTGTCTG 802
QY 781 CTGTACATCTACTCACTCAGATGAAGTAGCTGACACTGGCTGAGGCTGAGATTG 840
Db 803 CTGTACATCTACTCACTCAGATGAAGTAGCTGACACTGGCTGAGGCTGAGATTG 862
QY 841 CAGTGTCCGCTTGGAAGGGCGAGTCAGAGATGACTTCTGTGTGTATGACCGCTGTG 900
Db 863 CAGTGTCCGCTTGGAAGGGCGAGTCAGAGATGACTTCTGTGTGTATGACCGCTGTG 922
QY 901 TGAACATGATGGGTGGCAGGCCAATGATCTGTGATGATGGGGAGACTTAACCCACT 960
Db 923 TGAACATGATGGGTGGCAGGCCAATGATCTGTGATGATGGGGAGACTTAACCCACT 982
QY 961 GGGTTTAAAGATATCCAAACGTGTTTAAAGATCCGAGGCATTGTGGAAGAGAGCG 1020
Db 983 GGGTTTAAAGATATCCAAACGTGTTTAAAGATCCGAGGCATTGTGGAAGAGAGCG 1042
QY 1021 TGACTGGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAAGCTGTGTCCGGCCA 1080
Db 1043 TGACTGGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAAGCTGTGTCCGGCCA 1102
QY 1081 TGAACGTCAATGATTTCTGTACCAACAGAGATTGATACTTGTACTGTCCGAGAA 1140
Db 1103 TGAACGTCAATGATTTCTGTACCAACAGAGATTGATACTTGTACTGTCCGAGAA 1162
QY 1141 CCATTTTGGATGGCTGAAGAGGACACAGATGTGATGTTGGTGGAAACAAGTGTGG 1200
Db 1163 CCATTTTGGATGGCTGAAGAGGACACAGATGTGATGTTGGTGGAAACAAGTGTGG 1222
QY 1201 TGTGTGGCTATGGTGAAGTAGGCAAGGGCTGTGTCTCTCAAGCTCTTGAGCAAA 1260
Db 1223 TGTGTGGCTATGGTGAAGTAGGCAAGGGCTGTGTCTCTCAAGCTCTTGAGCAAA 1282
QY 1261 TTGTCTACATTAACGAATCGAACCCCATCTGTGCTGTGAGGCTGCATGATGGTTCA 1320
Db 1283 TTGTCTACATTAACGAATCGAACCCCATCTGTGCTGTGAGGCTGCATGATGGTTCA 1342
QY 1321 GGGTGGTAAAGCTAAATGAAGTCATCCGGCAAGTCGATGCTAATACTTCACAGGAA 1380
Db 1343 GGGTGGTAAAGCTAAATGAAGTCATCCGGCAAGTCGATGCTAATACTTCACAGGAA 1402
QY 1381 ATAGAATGTAGTGAACAGGGGAGCACTTGATCGCATGAATAAGTGTATCGTATGCA 1440
Db 1403 ATAGAATGTAGTGAACAGGGGAGCACTTGATCGCATGAATAAGTGTATCGTATGCA 1462
QY 1441 ATATGGGCCACTCCAACACAGAAATCGATGTGACACAGCCTCCGCACTCCGAGCTGACGT 1500
Db 1463 ATATGGGCCACTCCAACACAGAAATCGATGTGACACAGCCTCCGCACTCCGAGCTGACGT 1522
QY 1501 GGGAGCAGTACGTTCTCAGGTGAGCATGTCTGCGCAGATGGCAAAAGAGTTGTCC 1560
Db 1523 GGGAGCAGTACGTTCTCAGGTGAGCATGTCTGCGCAGATGGCAAAAGAGTTGTCC 1582
QY 1561 TCCTGGCAGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTTCCACCTTTGTCTGT 1620
Db 1583 TCCTGGCAGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTTCCACCTTTGTCTGT 1642

QY 1621 CCATCAGAGCCACACACAGGCTTTGGCACTGATAGACTCTATAATGCACCCGAGGGC 1680
Db 1643 CCATCAGAGCCACACACAGGCTTTGGCACTGATAGACTCTATAATGCACCCGAGGGC 1702
QY 1681 GATACAGCAGGATGTGACTTGTCTCTAAGAAATGATGAATACGTTGCCAGCTTGC 1740
Db 1703 GATACAGCAGGATGTGACTTGTCTCTAAGAAATGATGAATACGTTGCCAGCTTGC 1762
QY 1741 ATCTGCATCATTTGATGCCCACTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1800
Db 1763 ATCTGCATCATTTGATGCCCACTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1822
QY 1801 GACTCAACAAAAATGGGCCATTCAAACTTAATTATACAGATACTAATGACCATACTAC 1860
Db 1823 GACTCAACAAAAATGGGCCATTCAAACTTAATTATACAGATACTAATGACCATACTAC 1882
QY 1861 CAAGGACCAAGTCCACCTGAACCAACACACTCTAAGAAATATTTTAAAGATACTTTAT 1920
Db 1883 CAAGGACCAAGTCCACCTGAACCAACACACTCTAAGAAATATTTTAAAGATACTTTAT 1942
QY 1921 TTTCTTCTTACTCCTTTCCCTGTGATTTTTCCTATAATTTCAATCTTGTTTTCATC 1980
Db 1943 TTTCTTCTTACTCCTTTCCCTGTGATTTTTCCTATAATTTCAATCTTGTTTTCATC 2002
QY 1981 TCATTATCCAAAGTTCTGACAGACACACAGNACTTGCCTTCAAGGCTCTTGAAGAAATA 2040
Db 2003 TCATTATCCAAAGTTCTGACAGACACACAGNACTTGCCTTCAAGGCTCTTGAAGAAATA 2062
QY 2041 GAAGTTCAGGGTCCCTCAGCTCAGTCACTAAAGAGAGATTTTACTCCCCAGCCAGAAA 2100
Db 2063 GAAGTTCAGGGTCCCTCAGCTCAGTCACTAAAGAGAGATTTTACTCCCCAGCCAGAAA 2122
QY 2101 GGTGATTTCTCTTTACCATTTCTGGGAGCTTTAGTCTTAATTAGTACCTTATTAACA 2160
Db 2123 GGTGATTTCTT-TCTTTACCATTTCTGGGAGCTTTAGTCTTAATTAGTACCTTATTAACA 2181
QY 2161 GGAATGCTAAGGTACCTTCTCTGTGGAACAATCTGCAATGTCTAAATCGCCTTAAAGA 2220
Db 2182 GGAATGCTAAGGTACCTTCTCTGTGGAACAATCTGCAATGTCTAAATCGCCTTAAAGA 2241
QY 2221 GCCCATTTCTTAGCTGTGAATCAGTGTCTTTCACCTTCTCAGAGAGCAGGATGGT 2280
Db 2242 GCCCATTTCTTAGCTGTGAATCAGTGTCTTTCACCTTCTCAGAGAGCAGGATGGT 2301
QY 2281 ACCTACCAGGAGTAGGTAGTGTGGGTGTGATGTTAATTTCCCTTAGAAGTTCCA 2340
Db 2302 ACCTACCAGGAGTAGGTAGTGTGGGTGTGATGTTAATTTCCCTTAGAAGTTCCA 2361
QY 2341 AGCCCTGTTTCTGCTGAAGTGTGATGTCCAGTCCAGAGATGTGTATAATGAGCATGG 2400
Db 2362 AGCCCTGTTTCTGCTGAAGTGTGATGTCCAGTCCAGAGATGTGTATAATGAGCATGG 2421
QY 2401 CTTGTTAAGATCAGAGAGGCCCACTTGGATTATAGTATAGCCCTTCTCCACTCCACCA 2460
Db 2422 CTTGTTAAGATCAGAGAGGCCCACTTGGATTATAGTATAGCCCTTCTCCACTCCACCA 2481
QY 2461 GACTTGTCTATTTTTCGAGTTTAACTAGACTACACTCTATT 2503
Db 2482 GACTTGTCTATTTTTCGAGTTTAACTAGACTACACTCTATT 2524

RESULT 6
HSU82761 2258 bp mRNA linear PRI 17-DEC-2001
LOCUS
DEFINITION Homo sapiens S-adenosyl homocysteine hydrolase homolog (XPVKona)
ACCESSION U82761
VERSION U82761.1 GI:2852124
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 2258)
AUTHORS Cleaver,J.E., Afzal,V., Feeney,L., McDowell,M., Sadlinski,W.,
Volpe,J.P.G., Busch,D.B., Coleman,D.M., Ziffer,D.W., Yu,Y.,
Nagasawa,H. and Little,J.B.
TITLE Increased ultraviolet sensitivity and chromosomal instability
related to p53 function in the xeroderma pigmentosum variant
JOURNAL Cancer Res. 59 (5), 1102-1108 (1999)
MEDLINE 99168517
PUBMED 10070969
REFERENCE 2 (bases 1 to 2258)
AUTHORS Volpe,J.P.G., McDowell,M., Jostes,R.F., Afzal,V., Sadlinski,W.,
Trask,B.J., Legerstki,R. and Cleaver,J.E.
TITLE Complementation of chromosomal instability in the xeroderma
pigmentosum variant by a gene on human chromosome 1 with homology
to S-adenosyl homocysteine hydrolase
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 2258)
AUTHORS Volpe,J.P.G., McDowell,M. and Cleaver,J.E.
TITLE Direct Submission
JOURNAL Submitted (19-DEC-1996) Dermatology, UCSF, 3rd and Parnassus, Box
0750, San Francisco, CA 94143, USA
FEATURES
Source 1. 2258
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/cell_type="fibroblasts"
/tissue_type="skin"
1. 2258
/gene="XPvKona"
48. 1550
/note="xeroderma pigmentosum variant"
/codon_start=1
/product="S-adenosyl homocysteine hydrolase homolog"
/protein_id="AAC01960.1"
/db_xref="GI:2652125"
/translation="MATVTKAPKKOIOFADMOEFTKPKTKGRSLRSISQSSTDS
YSSAASYTDSDEVSPEKQNTSKSSNFCVKNIKQAEGRREIEAEDMSALIS
LKRAGEKPLAGAKIVGCTHTAQTAVLIETICALGACRSACNISTONEVAAL
AEAGVAFAMKGESEDDFWCIDRCVNDGMOANMILDDGDLTHWYKKYPNVFKI
RGIVESVTGVHRLYQLSKAGKLCVPMANVNDVTKQFNLYCCRESILDLKRTTD
VMFGKOVVCGYGEVKGKCAALKAIAIVYTEIDPICALQACMDGFRVVKLNEVI
RODVVITCTGNKNVNTREHIDRMKNSCIVCMGHSNTEIDVTSLRPELTWERVRSO
VDHVIMPDGKRVVLAEGRLNLSCSTVPTFVLSTATQALALIELYNAPEGRYKOD
VLLPKMDEYVASLHPSFDAHLTELTDQAKYLLGNKNGPFKNYYRY"
BASE COUNT 603 a 513 c 542 g 600 t
ORIGIN
Query Match 72.8%; Score 1866; DB 9; Length 2258;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2086; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 298 AGCTGAGCAGGCCAAGAGATCGAGACGCCGAGAGTACTCCTTCATGGCCACCGTCA 357
Db 1 AGCTGAGCAGGCCAAGAGATCGAGACGCCGAGAGTACTCCTTCATGGCCACCGTCA 60
QY 358 CCAAGCGGCCCAAGAGCAAAATCCAGTTTGTGATGACATGACAGAGTCCACCAATTC 417
Db 61 CCAAGCGGCCCAAGAGCAAAATCCAGTTTGTGATGACATGACAGAGTCCACCAATTC 120
QY 418 CCACCAAAACTGGCCGAGATCTTGTCTCGCTCGATCTCACAGTCCCTCCACTGACAGCT 477
Db 121 CCACCAAAACTGGCCGAGATCTTGTCTCGCTCGATCTCACAGTCCCTCCACTGACAGCT 180
QY 478 ACAGTTCAGCTGCATCTACACAGATAGCTCTGATGATGAGGTTTCCCGAGAGAAGC 537
Db 181 ACAGTTCAGCTGCATCTACACAGATAGCTCTGATGATGAGGTTTCCCGAGAGAAGC 240
QY 538 AGCAAAACCACTCCAGGGCAGACCAATTTCTGTGTAAGAACATCAAGCAGCAGAT 597
Db 241 AGCAAAACCACTCCAGGGCAGACCAATTTCTGTGTAAGAACATCAAGCAGCAGAT 300

QY 598 TTGAGCCCGGGAGATTGAGATTGCAGAGCAAGACATGTCTGCTGTATTCTACAGCA 657
Db 301 TTGAGCCCGGGAGATTGAGATTGCAGAGCAAGACATGTCTGCTGTATTCTACAGCA 360
QY 658 AACGTGCTCAGCGGGAGAGAGCCCTTGGCTGTGCTAAATAGTGGGCTGTACACACATCA 717
Db 361 AACGTGCTCAGCGGGAGAGAGCCCTTGGCTGTGCTAAATAGTGGGCTGTACACACATCA 420
QY 718 CAGCCCAAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGGGGCTCAGTGCCTGGT 777
Db 421 CAGCCCAAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGGGGCTCAGTGCCTGGT 480
QY 778 CTGCTGTACATCTACTCACTCAGAAATGAGTAGCTGCAGCAGCTGGCTGAGGCTGAG 837
Db 481 CTGCTGTACATCTACTCACTCAGAAATGAGTAGCTGCAGCAGCTGGCTGAGGCTGAG 540
QY 838 TTGCAGTGTTCGCTTGGAAGGGCGAGTCAGAAATGACTTCTGTGTGATTGACCGCT 897
Db 541 TTGCAGTGTTCGCTTGGAAGGGCGAGTCAGAAATGACTTCTGTGTGATTGACCGCT 600
QY 898 GTGTGAACATGATGGGTGGCAGGCCCAACATGATCTGTGATGATGGGAGACTTAACC 957
Db 601 GTGTGAACATGATGGGTGGCAGGCCCAACATGATCTGTGATGATGGGAGACTTAACC 660
QY 958 ACTGGTTTAAAGAGATATCCAAACGTGTTTAAAGATCCGAGGCAATGTGAAGAGA 1017
Db 661 ACTGGTTTAAAGAGATATCCAAACGTGTTTAAAGATCCGAGGCAATGTGAAGAGA 720
QY 1018 GCGTACTGTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTTCGG 1077
Db 721 GCGTACTGTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTTCGG 780
QY 1078 CCATGACGTCATGATTTCTGTACCAACAGAGATTGATACTGTACTGCTGCCGAG 1137
Db 781 CCATGACGTCATGATTTCTGTACCAACAGAGATTGATACTGTACTGCTGCCGAG 840
QY 1138 AATCCATTGATGATGGCTGGAAGAGACACAGATGTGATTTGTTGGGAACAAGTGG 1197
Db 841 AATCCATTGATGATGGCTGGAAGAGACACAGATGTGATTTGTTGGGAACAAGTGG 900
QY 1198 TGTGTGTGCTATGTGAGGTAGGCAAGGGCTGTGCTGCTCTCAAAAGCTCTTGAG 1257
Db 901 TGTGTGTGCTATGTGAGGTAGGCAAGGGCTGTGCTGCTCTCAAAAGCTCTTGAG 960
QY 1258 CAATGTCTACATTACCGAAATCGACCCCATCTGTGCTCTGCAGGCCCTGCATGATGG 1317
Db 961 CAATGTCTACATTACCGAAATCGACCCCATCTGTGCTCTGCAGGCCCTGCATGATGG 1020
QY 1318 TCAGGGTGTAAAGCTTAATGAAGTCATCCGGCAAGTCGATGCTGTAATACCTGCACAG 1377
Db 1021 TCAGGGTGTAAAGCTTAATGAAGTCATCCGGCAAGTCGATGCTGTAATACCTGCACAG 1080
QY 1378 GAAATAAGATGTAGTGACACAGGAGCACTTGGATGCGCATGAAAAACAGTTGATCGTAT 1437
Db 1081 GAAATAAGATGTAGTGACACAGGAGCACTTGGATGCGCATGAAAAACAGTTGATCGTAT 1140
QY 1438 GCAATATGGGCCACTCCACACAGAAATCGATGACCAAGCTCCGCACTCCGAGAGTGA 1497
Db 1141 GCAATATGGGCCACTCCACACAGAAATCGATGACCAAGCTCCGCACTCCGAGAGTGA 1200
QY 1498 CGTGGAGCGAGTACGTTCTCAGGTGACCATGTTCATCTGGCCAGATGGCAACAGAGTTG 1557
Db 1201 CGTGGAGCGAGTACGTTCTCAGGTGACCATGTTCATCTGGCCAGATGGCAACAGAGTTG 1260
QY 1558 TCCTCTGGCAGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTTCACACCTTTGTT 1617
Db 1261 TCCTCTGGCAGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTTCACACCTTTGTT 1320
QY 1618 TGTTCATCAGCCACACACAGAGGCTTTGGCACTGATAGAACTCTATATGACCCGAGG 1677
Db 1321 TGTTCATCAGCCACACACAGAGGCTTTGGCACTGATAGAACTCTATATGACCCGAGG 1380

QY	1678	GCGGATACAAAGCAGAGTGTGTACTTGCTTCCTTAAGAAAAATGGATGAATACGTTGCCAGCT	1737
Db	1381	GGCGATACAAGCAGAGTGTGTACTTGCTTCCTTAAGAAAAATGGATGAATACGTTGCCAGCT	1440
QY	1738	TGCATCTGCCATCATTTGATGCCCCACCCTTACAGAGCTGACAGATGACCAAGCAAATAATAC	1797
Db	1441	TGCATCTGCCATCATTTGATGCCCCACCCTTACAGAGCTGACAGATGACCAAGCAAATAATAC	1500
QY	1798	TGGGACTCAACAAAATGGGCCATTCAAACCTAATTATTACAGATACTAATGGACCATAAC	1857
Db	1501	TGGGACTCAACAAAATGGGCCATTCAAACCTAATTATTACAGATACTAATGGACCATAAC	1560
QY	1858	TACCAAGGACCAGTCCACCTGTAACACACACACTCTAAAGAAAATTTTTTAAGATAACTTT	1917
Db	1561	TACCAAGGACCAGTCCACCTGTAACACACACACTCTAAAGAAAATTTTTTAAGATAACTTT	1620
QY	1918	TATTTCTCTTACTCCCTTCCCTCTGATTTTTTTTCCCTAATTTCAATTCCTGTGTTTTTC	1977
Db	1621	TATTTCTCTTACTCCCTTCCCTCTGATTTTTTTTCCCTAATTTCAATTCCTGTGTTTTTC	1680
QY	1978	ATCTCATTTATCCAAGTCTGTGAGACCACACAGGAACCTTGCTTCATGGCTCTTAGATGAA	2037
Db	1681	ATCTCATTTATCCAAGTCTGTGAGACCACACAGGAACCTTGCTTCATGGCTCTTAGATGAA	1740
QY	2038	ATAGAAGTTCAGGGTCCCCTCACCTCTAGTCACCTAAGAAGAGATTTTACTCCCCCAGCCAG	2097
Db	1741	ATAGAAGTTCAGGGTCCCCTCACCTCTAGTCACCTAAGAAGAGATTTTACTCCCCCAGCCAG	1800
QY	2098	AAAGGTGATTTCTCTCTTTACCATTCTCTGGGGACTTGTAGTCTTAATTAAGGTACCTTATTA	2157
Db	1801	AAAGGTGATTTCT - TCTTTACCATTCTCTGGGGACTTGTAGTCTTAATTAAGGTACCTTATTA	1859
QY	2158	ACAGGAATGCTPAAGGTACCTTCTCTGTGGAACAATCTGCAATGTCTAAATCGCCTTAAA	2217
Db	1860	ACAGGAATGCTPAAGGTACCTTCTCTGTGGAACAATCTGCAATGTCTAAATCGCCTTAAA	1919
QY	2218	AGAGCCCATTTCTTAGCTGTGGAATACTAGTGTCTCTTACCTCTTACAGAGAAGCAGGAT	2277
Db	1920	AGAGCCCATTTCTTAGCTGTGGAATACTAGTGTCTCTTACCTCTTACAGAGAAGCAGGAT	1979
QY	2278	GGTACCTACCCGGCAGAGTAGTTAGATGTGGGTGTGTGATGTCCAGTTAATTTCCCTTAGAAGT	2337
Db	1980	GGTACCTACCCGGCAGAGTAGTTAGATGTGGGTGTGTGATGTCCAGTTAATTTCCCTTAGAAGT	2039
QY	2338	CCAAGCCCTGTTTCTCTGCGTAAAGGTGGTATGTCCAGTTCAGAGATGTG 2386	
Db	2040	CCAAGCCCTGTTTCTCTGCGTAAAGGTGGTATGTCCAGTTCAGAGATGTG 2088	

RESULT 7	
MSM800298	
LOCUS	2510 bp mRNA linear PRI 18-FEB-2000
DEFINITION	Homo sapiens mRNA; CDNA DKFZp564A1523 (from clone DKFZp564A1523);
ACCESSION	AL049954
VERSION	AL049954.1 GI:4884203
KEYWORDS	
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
19	(pages 1 to 2510)	Wambutt, R., Heudner, D., Mewes, H.W., Gassenhuber, J. and Wiemann, S.	Direct Submission	
		Submitted (15-MAY-1999)	MIPS, Am Klopferspitz 18a, D-82152	
		Marlinsried, GERMANY		
		Clone from S. Wiemann, Molecular Genome Analysis, German Cancer		

This clone (DKR2p564A1523) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further

information about the clone and the sequencing project is available at <http://www.mips.biochem.mpg.de/proj/cDNA/>.

Location/Qualifiers

1: .2510

FEATURES

source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="1, 15.55 cR from DIS248"
/clone="DKFZp564A1523"
/tissue_type="brain"
/clone_1lb="564 (synonym: hfbz2). Vector pAMPl; host
xl-2blue; sites NotI + SalI"
/dev_stage="fetal"
1. .1796
/gene="DKFZp564A1523"
<1. .1796

```

/gene="DnAdp204n123"
 /note="S-adenosyl homocysteine hydrolase homolog XPvkona
 splice variant"
 /codon_start=3
 /product="hypothetical protein"
 /protein_id="CAB43223.1"
 /db_xref="GI:4884204"
 /translation="GLITVELDSSCVCIILLLRHGGSRGKICAEISOPALPLPOE
 TQAPSGESENLRNLRQGRAESSQNLSLRRLVSAQNGRSDWRMLNQLMPEIDSL
 ISCFLOIQFADDMQEFTEKPTKTRGRSLRSLSQSSTDSYSSAASITDSDDEVSPRE
 KQQTNSKGSNFCVKNKIQAEFRGEIEIAEDMSALISLRRAOGEKPLAGAKIVGC
 THITAOATAVLIETICALGAOCRMSACNITSONEVAALAEAGVAVFAMGSESEDEFW
 WCIDRCVNDGMOANMILDDGGDLTHWYKYPNVFEKKIRGIVEESVTGYHRLYQLSK
 AGKLCVPAMVNDSVTKQKFDNLYCCRESIDLGLKRTDVMFEGKQVYVCGYGEVKGK
 CCAALKALGAIYVITEIDPICALQACMDGFYVVKLNEVIRQDVVITCTGNKNVVTRE
 HLRMKNSCIYCNMGHSNTEIDVTSLRTPELTWERVRSQVDHVITWPDGRVYLLAEGR
 LLNLSCTVPFEVLSITATQALALIELYNADEGRYKQDVYLLPKMDEYVASLHLPS
 FDAHLELTDQAKYLGINKNGPFKPNYYRY"

polyA site	2490		
BASE COUNT	664 a	556 c	618 g
ORIGIN			672 t

Query Match	72.48;	Score 1856;	DB 9;	Length 2510;	
Best Local Similarity	99.88;	Pred. No. 0;			
Matches 2126; Conservative	0;	Mismatches 3;	Indels 1;	Gaps 1;	

OY	374	GCAAAATCCAGTTTGCTGATGACATGCAGAGTTCACCAAATTCCCACCACCAAAACTGGCCG	433
Db	323	GCAAAATCCAGTTTGCTGATGACATGCAGAGTTCACCAAATTCCCACCACCAAAACTGGCCG	382
OY	434	AAGATCTTTGTCTCGCTCGATCTCACAGTCTCCACTGACAGCTACAGTTCAGCTGCATC	493
Db	383	AAGATCTTTGTCTCGCTCGATCTCACAGTCTCCACTGACAGCTACAGTTCAGCTGCATC	442
OY	494	CTACACAGATAGCTCTGATGATGAGTTTCTCCCCGAGAGAAGCAGCAAAACAACCTCCA	553
Db	443	CTACACAGATAGCTCTGATGATGAGTTTCTCCCCGAGAGAAGCAGCAAAACAACCTCCA	502
OY	554	GGGCGACGACCATTTCTGTGTGAAGAACAATCAAGCAGGACAGATTTGACGCCGGGAGAT	613
Db	503	GGGCGACGACCATTTCTGTGTGAAGAACAATCAAGCAGGACAGATTTGACGCCGGGAGAT	562
OY	614	TGAGATTGCAGAGCAAGACATGTCTGCTTGATTTCACTCAGGAACGTGCTCAGGGGA	673
Db	563	TGAGATTGCAGAGCAAGACATGTCTGCTTGATTTCACTCAGGAACGTGCTCAGGGGA	622
OY	674	GAAGCCCTTGGCTGGTGCTAAAATAGTGGGCTGTACACACATCACAGCCACAGACGCCGT	733
Db	623	GAAGCCCTTGGCTGGTGCTAAAATAGTGGGCTGTACACACATCACAGCCACAGACGCCGT	682
OY	734	GTTGATTGAGACACTCTGTGCCCTGGGGGCTCAGTGCCGCTGGTCTGCTTGTAACATCTA	793
Db	683	GTTGATTGAGACACTCTGTGCCCTGGGGGCTCAGTGCCGCTGGTCTGCTTGTAACATCTA	742
OY	794	CTCAACTCAGAAATGAAGTAGCTGCAGACACTGGCTGAGGCTGGAGTTGCAGTGTTCGCTTG	853
Db	743	CTCAACTCAGAAATGAAGTAGCTGCAGACACTGGCTGAGGCTGGGGTTGCAGTGTTCGCTTG	802

QY 854 GAAGGCGAGTCAGAGATGACTTCTGGTGTATTTGACCGCTGTGTGAACATGGATGG 913
|||||
Db 803 GAAGGCGAGTCAGAGATGACTTCTGGTGTATTTGACCGCTGTGTGAACATGGATGG 862
QY 914 GTGGCAGGCCAACATGATCCTGGATGATGGGGGAGACTTAACCCACTGGGTTATAGAA 973
|||||
Db 863 GTGGCAGGCCAACATGATCCTGGATGATGGGGGAGACTTAACCCACTGGGTTATAGAA 922
QY 974 GTATCCAAACGTGTTAAGAGATCCGAGCATTTGTGAAGAGAGCGGTGACTGTGTCA 1033
|||||
Db 923 GTATCCAAACGTGTTAAGAGATCCGAGCATTTGTGAAGAGAGCGGTGACTGTGTCA 982
QY 1034 CAGGCTGATCAGCTCTCCAAAGCTGGAGCTGTGTCCGGCCATGAACGTCAATGA 1093
|||||
Db 983 CAGGCTGATCAGCTCTCCAAAGCTGGAGCTGTGTCCGGCCATGAACGTCAATGA 1042
QY 1094 TTCTGTACCAACAGAGATTGTATTAATCTGTACTGCTGCCGAGATCCATTTGATGG 1153
|||||
Db 1043 TTCTGTACCAACAGAGATTGTATTAATCTGTACTGCTGCCGAGATCCATTTGATGG 1102
QY 1154 CCTGAGAGAGACACAGATGTGATGTTGGTGGGAACAAGTGTGTGTGGCTATGG 1213
|||||
Db 1103 CCTGAGAGAGACACAGATGTGATGTTGGTGGGAACAAGTGTGTGTGGCTATGG 1162
QY 1214 TGAGGTAGGCCAAGGGCTGTGTGTCTCAAAAGCTCTTGAGCAATGTCTACATTAC 1273
|||||
Db 1163 TGAGGTAGGCCAAGGGCTGTGTGTCTCAAAAGCTCTTGAGCAATGTCTACATTAC 1222
QY 1274 CGAAATCGACCCCATCTGTGTCTGTGACGGCTGCATGGATGGGTTCAAGGTTGTAAGCT 1333
|||||
Db 1223 CGAAATCGACCCCATCTGTGTCTGTGACGGCTGCATGGATGGGTTCAAGGTTGTAAGCT 1282
QY 1334 AAATGAGTCAATCCGGCAGTCGATGCTGAATTAATCTTGACAGAGAAATGAATGTAGT 1393
|||||
Db 1283 AAATGAGTCAATCCGGCAGTCGATGCTGAATTAATCTTGACAGAGAAATGAATGTAGT 1342
QY 1394 GACACGGGAGCACTTGATCGCATGAAAAACAGTTGTATCGTATGCAATATGGGCCACTC 1453
|||||
Db 1343 GACACGGGAGCACTTGATCGCATGAAAAACAGTTGTATCGTATGCAATATGGGCCACTC 1402
QY 1454 CAACACAGAAATCGATGTGACACAGCCTCCGCACTCCGAGCTGACGTGGAGCCGACTACG 1513
|||||
Db 1403 CAACACAGAAATCGATGTGACACAGCCTCCGCACTCCGAGCTGACGTGGAGCCGACTACG 1462
QY 1514 TTCTCAGGTGGACCATGTCTATGCGCCAGATGGCAACAGAGTTGCTCTGCGCAGAGGG 1573
|||||
Db 1463 TTCTCAGGTGGACCATGTCTATGCGCCAGATGGCAACAGAGTTGCTCTGCGCAGAGGG 1522
QY 1574 TCGTCTACTCAATTTGAGCTGCTCCACAGTTCCCACTTTGTTCTGTCCATCAGCCAC 1633
|||||
Db 1523 TCGTCTACTCAATTTGAGCTGCTCCACAGTTCCCACTTTGTTCTGTCCATCAGCCAC 1582
QY 1634 AACACAGGCTTTGGCACTGATAGAACTCTATATGCAACCCGAGGGGCGATACAGCAGGA 1693
|||||
Db 1583 AACACAGGCTTTGGCACTGATAGAACTCTATATGCAACCCGAGGGGCGATACAGCAGGA 1642
QY 1694 TGTGTACTTGTCTCTTAAGAAATGATGAATACGTTGCCAGCTTGCAATCTGCCATCAT 1753
|||||
Db 1643 TGTGTACTTGTCTCTTAAGAAATGATGAATACGTTGCCAGCTTGCAATCTGCCATCAT 1702
QY 1754 TGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGAGCTCAACAAAA 1813
|||||
Db 1703 TGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGAGCTCAACAAAA 1762
QY 1814 TGGGCCATTCAAACCTAATTATTACAGATTAATGACCATACTACCAAGCAGCAGTCC 1873
|||||
Db 1763 TGGGCCATTCAAACCTAATTATTACAGATTAATGACCATACTACCAAGCAGCAGTCC 1822
QY 1874 ACCTGACACACACACTCTAAAGAAATATTTTAAGATAACTTTATTTCTCTACTC 1933
|||||
Db 1823 ACCTGACACACACACTCTAAAGAAATATTTTAAGATAACTTTATTTCTCTACTC 1882
QY 1934 CTTTCTCTTGTATTTTTCCTATAATTCATCTCTTTTTCATCTCATATCCAAAGT 1993

Db 1883 CTTTCTCTTGTATTTTTCCTATAATTCATCTCTTTTTCATCTCATATCCAAAGT 1942
QY 1994 TCTGACACACACAGGAACCTGCTTCATGGCTCTTTAGATGAATAGAGTTCAGGGTC 2053
|||||
Db 1943 TCTGACACACACAGGAACCTGCTTCATGGCTCTTTAGATGAATAGAGTTCAGGGTC 2002
QY 2054 CCTCACTCTAGTCACTAAAGAGATTTTACCTCCCAAGCCCAAGAAAGTGATCTCTC 2113
|||||
Db 2003 CCTCACTCTAGTCACTAAAGAGATTTTACCTCCCAAGCCCAAGAAAGTGATCTCTC 2061
QY 2114 TTACCATTTCTGGGACTTTAGTCTTAATTAGTACCTTATTACAGGAATGCTAAG 2173
|||||
Db 2062 TTACCATTTCTGGGACTTTAGTCTTAATTAGTACCTTATTACAGGAATGCTAAG 2121
QY 2174 TACCTTCTGTGGAACAATCTGCAATGCTTAATGCGCTTAAAGAGCCCATTTCTAG 2233
|||||
Db 2122 TACCTTCTGTGGAACAATCTGCAATGCTTAATGCGCTTAAAGAGCCCATTTCTAG 2181
QY 2234 CTGCTGAATCAGTCTCTTTTCACTTCTTACAGAGACAGGGATGTACCTACCCGGCAG 2293
|||||
Db 2182 CTGCTGAATCAGTCTCTTTTCACTTCTTACAGAGACAGGGATGTACCTACCCGGCAG 2241
QY 2294 GTAGGTTAGTGGGCTGTGTCATGTTAATTTCCCTTAGAAGTTCCAAAGCCCTGTTCT 2353
|||||
Db 2242 GTAGGTTAGTGGGCTGTGTCATGTTAATTTCCCTTAGAAGTTCCAAAGCCCTGTTCT 2301
QY 2354 GCGTAAAGTGGTATGTCCAGTTCCAGAGATGTATATAGACATGGCTGTTAAGATCA 2413
|||||
Db 2302 GCGTAAAGTGGTATGTCCAGTTCCAGAGATGTATATAGACATGGCTGTTAAGATCA 2361
QY 2414 GGAGGCCCATTTGATTTATAGTATAGCCCTTCCCTCCACTCCCAACAGACTGCTCATTT 2473
|||||
Db 2362 GGAGGCCCATTTGATTTATAGTATAGCCCTTCCCTCCACTCCCAACAGACTGCTCATTT 2421
QY 2474 TTGAGTTTAACTAGACTACACTTAAT 2503
|||||
Db 2422 TTGAGTTTAACTAGACTACACTTAAT 2451

RESULT 8
AF035319 1369 bp mRNA linear PRI 07-JUL-2000
LOCUS AF035319
DEFINITION Homo sapiens clone 23931 mRNA, partial cds.
ACCESSION AF035319
VERSION AF035319.1 GI:2661082
KEYWORDS FLI_CDNA.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 1369)
AUTHORS Andersson,B., Wentland,M.A., Ricafrente,J.Y., Liu,W. and gibbs,R.A.
TITLE A 'double adaptor' method for improved shotgun library construction
JOURNAL Anal. Biochem. 236 (1), 107-113 (1996)
MEDLINE 96207227
PUBMED 8619474
REFERENCE 2 (bases 1 to 1369)
AUTHORS Yu,W., Andersson,B., Worley,K.C., Muzny,D.M., Ding,Y., Liu,W.,
Ricafrente,J.Y., Wentland,M.A., Lennon,G. and gibbs,R.A.
TITLE Large-scale concatenation cDNA sequencing
JOURNAL Genome Res. 7 (4), 353-358 (1997)
MEDLINE 97264341
PUBMED 9110174
REFERENCE 3 (bases 1 to 1369)
AUTHORS Yu,W., Sarginson,J. and gibbs,R.A.
TITLE Direct Submission
JOURNAL Submitted (20-NOV-1997) Molecular and Human Genetics, Baylor
College of Medicine, One Baylor Plaza S930, Houston, TX 77030, USA
FEATURES
source
1. 1369
/organism="Homo sapiens"
/db_xref="taxon:9606"

/clone="23931"
/sex="female"
/tissue_type="brain"
/clone_id="1N1B"
/dev_stage="infant"
/note="This clone is similar to Drosophila M. Abd-B gene
encoded by GenBank Accession Number U07835. The I.M.A.G.E.
Consortium clone ID number is 23931"
<1.657

/note="similar to S-adenosylhomocysteine hydrolase"
/codon_start=1
/protein_id="AAB8189.1"
/db_xref="GI:2661083"
/translation="PRAMCGYGEVKGCCAAKALGAIYITEIDPICALQACMDGFR
VVKLNEYROYDVITCTGNKNVTRHLDKMKNSCIYCNMHSNTEIDVTSLRPEL
TWERVRSOVDHIVPDGKRVLLAEGRLNLSCTVPEVLSITATTOALALIELYNA
PEGRIKQDVYLLPKMDEVASLHLPFDALHLELTDQAKYLLGLNKNKPFKPNYRY

BASE COUNT 370 a 314 c 286 g 399 t
ORIGIN

Query Match 42.1%; Score 1080; DB 9; Length 1369;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1300; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1201 TGTGTGCTATGCTAGGTAGGCAAGGGCTGCTGCTGCTCTCAAGCTCTTGAGCA 1260
DB 11 TGTGTGCTATGCTAGGTAGGCAAGGGCTGCTGCTGCTCTCAAGCTCTTGAGCA 70
QY 1261 TTGCTCTACATTACCGAATCGACCCCATCTGTGCTCTGACAGGCTGATGGGTCA 1320
DB 71 TTGCTCTACATTACCGAATCGACCCCATCTGTGCTCTGACAGGCTGATGGGTCA 130
QY 1321 GGGTGGTAAAGCTAAATGAAGTCATCCGGCAAGTCGATGCTAAATTAAGTTCACAGGAA 1380
DB 131 GGGTGGTAAAGCTAAATGAAGTCATCCGGCAAGTCGATGCTAAATTAAGTTCACAGGAA 190
QY 1381 ATAAGATGTAGTGACACGGGAGCAGCTTGATCGCATGAAAAACAGTTGATCGTATGCA 1440
DB 191 ATAAGATGTAGTGACACGGGAGCAGCTTGATCGCATGAAAAACAGTTGATCGTATGCA 250
QY 1441 ATATGGGCCACTCCACACAGAAATCGATGTGACCAAGCCTCCGCACTCCGAGCTGACGT 1500
DB 251 ATATGGGCCACTCCACACAGAAATCGATGTGACCAAGCCTCCGCACTCCGAGCTGACGT 310
QY 1501 GGGAGCGAGTACGTTCTCAGGTGACCATGTCATCTGGCCAGATGGCAACGAGTTGTC 1560
DB 311 GGGAGCGAGTACGTTCTCAGGTGACCATGTCATCTGGCCAGATGGCAACGAGTTGTC 370
QY 1561 TCCTGGCAGAGGCTGCTACTCAATTTGAGCTGCTCCACAGTTCCACCTTTGTTCTGT 1620
DB 371 TCCTGGCAGAGGCTGCTACTCAATTTGAGCTGCTCCACAGTTCCACCTTTGTTCTGT 430
QY 1621 CCATCACAGCCACACACAGGCTTTGGCACTGATAGAACTCTATAATGCAACCCGAGGGGC 1680
DB 431 CCATCACAGCCACACACAGGCTTTGGCACTGATAGAACTCTATAATGCAACCCGAGGGGC 490
QY 1681 GATACACAGCAGATGTGTACTGCTTCTTAAGAAAAATGATGAATACGTTGCCAGCTTGC 1740
DB 491 GATACACAGCAGATGTGTACTGCTTCTTAAGAAAAATGATGAATACGTTGCCAGCTTGC 550
QY 1741 ATGAGCCATCATTTGATGCCCACTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1800
DB 551 ATGAGCCATCATTTGATGCCCACTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 610
QY 1801 GACTCAACAAAAATGGGCACTCAAACTTAATATATACAGATACATAATGACCACTACTAC 1860
DB 611 GACTCAACAAAAATGGGCACTCAAACTTAATATATACAGATACATAATGACCACTACTAC 670
QY 1861 CAAGGACCACTCCACCTGAACACACACTCTAAAGAAATATTTTAAAGATTAAGTCTTAT 1920
DB 671 CAAGGACCACTCCACCTGAACACACACTCTAAAGAAATATTTTAAAGATTAAGTCTTAT 730

QY 1921 TTCTCTTACTCCTCTTCCCTCTGATTTTTCCTATAATTTCAATCTTGTGTTTTCATC 1980
DB 731 TTCTCTTACTCCTCTTCCCTCTGATTTTTCCTATAATTTCAATCTTGTGTTTTCATC 790
QY 1981 TCATTAACCAAGTCTCTGACAGCCACACAGAACTGCTTCATGGCTCTTAGATGAATA 2040
DB 791 TCATTAACCAAGTCTCTGACAGCCACACAGAACTGCTTCATGGCTCTTAGATGAATA 850
QY 2041 GAAGTTCAGGGTCCCTCACTCTAGTCACTAAGAAGATTTTACTCCCGCAGCCAGAAA 2100
DB 851 GAAGTTCAGGGTCCCTCACTCTAGTCACTAAGAAGATTTTACTCCCGCAGCCAGAAA 910
QY 2101 GGTGATTTCTTCTTTTACCATTTCTGGGACTTTAGTCTTAATAGGTACCTTAATA 2160
DB 911 GGTGATTTCTTCTTTTACCATTTCTGGGACTTTAGTCTTAATAGGTACCTTAATA 969
QY 2161 GGAATGCTAAGGTACCTCTCTGTGGAACAATCTGCAATGTCTAAATCGCCTTAAGA 2220
DB 970 GGAATGCTAAGGTACCTCTCTGTGGAACAATCTGCAATGTCTAAATCGCCTTAAGA 1029
QY 2221 GCCCATTTCTAGCTGCTGAAATCAGTGTCTTCACTTCTTCAGAGAGAGGATGGT 2280
DB 1030 GCCCATTTCTAGCTGCTGAAATCAGTGTCTTCACTTCTTCAGAGAGAGGATGGT 1089
QY 2281 ACCATCCGCGCAGGTAGGTAGATGTGGGTGTCATGTTAATTTCCCTTAGAGTTCCA 2340
DB 1090 ACCATCCGCGCAGGTAGGTAGATGTGGGTGTCATGTTAATTTCCCTTAGAGTTCCA 1149
QY 2341 AGCCCTGTTTCTGCGTAAAGGTGATGTCAGTTCAGAGATGTATATAAGACATGG 2400
DB 1150 AGCCCTGTTTCTGCGTAAAGGTGATGTCAGTTCAGAGATGTATATAAGACATGG 1209
QY 2401 CTGTAGATCAGGAGGCCCACTTGATTTATAGATAGCCCTTCCCTCCACCCACCA 2460
DB 1210 CTGTAGATCAGGAGGCCCACTTGATTTATAGATAGCCCTTCCCTCCACCCACCA 1269
QY 2461 GACTGCTCATTTTTCGAGTTTTCATCTAGTCACTCACTTAT 2503
DB 1270 GACTGCTCATTTTTCGAGTTTTCATCTAGTCACTCACTTAT 1312

RESULT 9
BC003631 1323 bp mRNA 1linear PRI 12-JUL-2001
LOCUS
DEFINITION Homo sapiens, Similar to S-adenosylhomocysteine hydrolase-like 1,
clone IMAGE:3010755, mRNA, partial cds.
ACCESSION BC003631
VERSION BC003631.1 GI:13177702
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
REFERENCE
AUTHORS
TITLES
JOURNAL
Submitted (26-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
COMMENT NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
contact: amadane@systemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
Greene, Mark Ketteman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 3 Row: b Column: 6.

Location/Qualifiers
1..1323

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3010755"
/tissue_type="Muscle, rhabdomyosarcoma"
/clone_id="NH_MGC_17"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
1..611
/codon_start=3
/product="Similar to S-adenosylhomocysteine hydrolase-like
1"

CDS

/protein_id="AAH03631.1"
/db_xref="GI:13177703"
/translation="HEEALGAIIVITEIDPICALQACMDGERVVKLNEYIROVDVIT
CTGNKNVTRHELDRLMKNSCIYCNMHSNTEIDVTSIRPELWERYRSQVDHVIWD
GKRVLLAEGRLNLSCSTVPTFVLSITATQALALIELYNAPGRYKQDYLPLPKM
DEVASLHLPSEDAHLELDDQAKYLGINKNGPKPNYRY"

BASE COUNT 365 a 303 c 271 g 384 t

ORIGIN

Query Match 40.48; Score 1035; DB 9; Length 1323;
Best Local Similarity 99.88; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

OY 1246 AAGCTCTGGAGCAATGTCTACATTAACGAAATCGACCCCATCTGTCTGACGGCT 1305
|||
Db 10 AAGCTCTGGAGCAATGTCTACATTAACGAAATCGACCCCATCTGTCTGACGGCT 69
OY 1306 GCATGGATGGCTCAGGGTGTAAAGCTAAATGAAGTATCCGGCAAGTCGATGCTTAA 1365
|||
Db 70 GCATGGATGGCTCAGGGTGTAAAGCTAAATGAAGTATCCGGCAAGTCGATGCTTAA 129
OY 1366 TAAGTGCACAGGAATAAGATGTAGTACACGGGAGCACTTGATCGCATGAANAACA 1425
|||
Db 130 TAAGTGCACAGGAATAAGATGTAGTACACGGGAGCACTTGATCGCATGAANAACA 189
OY 1426 GTTGTATCGTATGCAATATGGGCCACTCCAACACAGAAATCGATGTGACCAAGCTCCGCA 1485
|||
Db 190 GTTGTATCGTATGCAATATGGGCCACTCCAACACAGAAATCGATGTGACCAAGCTCCGCA 249
OY 1486 CTCGGAGCTGACGTGGGAGCGAGTACGTTCTCAGGTGACCATGTCTGCGCCAGATG 1545
|||
Db 250 CTCGGAGCTGACGTGGGAGCGAGTACGTTCTCAGGTGACCATGTCTGCGCCAGATG 309
OY 1546 GCAACGAGTGTCTCTCTGCGAGAGGGTGTCTACTCAATTGAGCTGCTCCACAGTTC 1605
|||
Db 310 GCAACGAGTGTCTCTCTGCGAGAGGGTGTCTACTCAATTGAGCTGCTCCACAGTTC 369
OY 1606 CCACCTTGTCTGTCCATCACAGCCACACACAGGCTTTGGCACTGATAGAACTCTATA 1665
|||
Db 370 CCACCTTGTCTGTCCATCACAGCCACACACAGGCTTTGGCACTGATAGAACTCTATA 429
OY 1666 ATGCACCCGAGGGCGGATACAGCAGGATGTGTACTTGTCTCTTAAGAAATGGATGAAT 1725
|||
Db 430 ATGCACCCGAGGGCGGATACAGCAGGATGTGTACTTGTCTCTTAAGAAATGGATGAAT 489
OY 1726 ACGTGGCCAGCTTGCACTTGCCATCATTTGATGCCCACTTACAGAGCTGACAGATGACC 1785
|||
Db 490 ACGTGGCCAGCTTGCACTTGCCATCATTTGATGCCCACTTACAGAGCTGACAGATGACC 549
OY 1786 AAGCAAAATATCTGGGACTCAACAAAATGGGCAATTCAACTAATTATTACAGATACT 1845
|||
Db 550 AAGCAAAATATCTGGGACTCAACAAAATGGGCAATTCAACTAATTATTACAGATACT 609
OY 1846 AATGACCACTACTACAGGACAGCAGTCCACCTGAACCACTCTAAAGAAATATTATT 1905
|||
Db 610 AATGACCACTACTACAGGACAGCAGTCCACCTGAACCACTCTCTAAAGAAATATTATT 669
OY 1906 TAAGTAACTTTTATTTCTTCTACTCCTTCTCTGATTTTTTCTCTATATTTCAT 1965

|||||
Db 670 TAAGTAACTTTTATTTCTTCTACTCCTCTCTGATTTTTTCTCTATATTTCAT 729
OY 1966 TCTTGTTTTTCATCTCATATATCCAGTCTGCGAGCCACACAGGAACTTGCTTCATGGC 2025
|||
Db 730 TCTTGTTTTTCATCTCATATATCCAGTCTGCGAGCCACACAGGAACTTGCTTCATGGC 789
OY 2026 TCTTGTAGAAATAGAGTTCAGGGTCCCTCACTCTAGTCACTAAAGAAGATTTTACT 2085
|||
Db 790 TCTTGTAGAAATAGAGTTCAGGGTCCCTCACTCTAGTCACTAAAGAAGATTTTACT 849
OY 2086 CCCCAGCCGAGAAAGGTGATCTCTCTTTTACATTTCTGGGCACTTTAGTCTTAATTA 2145
|||
Db 850 CTCCAGCCGAGAAAGGTGATCTCTCTTTTACATTTCTGGGCACTTTAGTCTTAATTA 908
OY 2146 GGTACCTTATTAACAGGAATGCTAAGTACCTTCTGTGGAACAATCTGCAATGTCTA 2205
|||
Db 909 GGTACCTTATTAACAGGAATGCTAAGTACCTTCTGTGGAACAATCTGCAATGTCTA 968
OY 2206 AATGCGCTTAAAGAGCCCATTTCTTAGCTGCTGAATCAGTCTCTTCACTTCTTCA 2265
|||
Db 969 AATGCGCTTAAAGAGCCCATTTCTTAGCTGCTGAATCAGTCTCTTCACTTCTTCA 1028
OY 2266 AGAAGCAGGAGTGTACCTACCCGGCAGGTAGGTAGATGTGGGTGTCATGTTAATT 2325
|||
Db 1029 AGAAGCAGGAGTGTACCTACCCGGCAGGTAGGTAGATGTGGGTGTCATGTTAATT 1088
OY 2326 CCCTTAGAAGTTCAGACCCCTGTTCTCTGCGTAAAGGTGATGTCCAGTTCAGAGATGT 2385
|||
Db 1089 CCCTTAGAAGTTCAGACCCCTGTTCTCTGCGTAAAGGTGATGTCCAGTTCAGAGATGT 1148
OY 2386 GTATATGACATGGCTGTTAAGATCAGAGAGCCCACTTGATTTATAGTATAGCCCTT 2445
|||
Db 1149 GTATATGACATGGCTGTTAAGATCAGAGAGCCCACTTGATTTATAGTATAGCCCTT 1208
OY 2446 CCTCCACTCCACACAGACTTGCTCATTTTTCGAGTTTCTTACTAGACTACACTCTAT 2503
|||
Db 1209 CCTCCACTCCACACAGACTTGCTCATTTTTCGAGTTTCTTACTAGACTACACTCTAT 1266

RESULT 10
LOCUS AR068569 2226 bp DNA linear PAT 29-SEP-1999
AR068569
DEFINITION Sequence 2 from patent US 5854023.
AR068569
ACCESSION AR068569
VERSION AR068569.1 GI:6000776
KEYWORDS
SOURCE
ORGANISM
unknown.
unknown.
unclassified.

REFERENCE 1 (bases 1 to 2226)
AUTHORS Hillman,J.L., Corley,N.C., Lal,P. and Shah,P.
TITLE Polynucleotides encoding human S-adenosyl-5-homocysteine hydrolase
derived from bladder
JOURNAL Patent: US 5854023-A 2 29-DEC-1998;
FEATURES
LOCATION/Qualifiers
source 1..2226

BASE COUNT 649 a 455 c 435 g 675 t 12 others
ORIGIN

Query Match 21.68; Score 553; DB 6; Length 2226;
Best Local Similarity 99.68; Pred. No. 7.7e-311;
Matches 773; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

OY 1728 GTTGCAGCTTGACATCTGCATCATTTGATGCCCACTTACAGAGCTGACAGATGACCA 1787
|||
Db 1 GTTGCAGCTTGACATCTGCATCATTTGATGCCCACTTACAGAGCTGACAGATGACCA 60
OY 1788 GCAAAATATCTGGGACTCAACAAAATGGGCAATTCAAACCTAATTATTACAGATCTAA 1847
|||
Db 61 GCAAAATATCTGGGACTCAACAAAATGGGCAATTCAAACCTAATTATTACAGATCTAA 120

OY	1848	TGGACCATCTACTACCAAGGACCAGTCCACCTCGAACCCACACACTCTAAAGAAATATTTT	TA	1907
Db	121	TGGACCATCTACTACCAAGGACCAGTCCACCTCGAACCCACACACTCTAAAGAAATATTTT	TA	180
OY	1908	AGATAACTTTTTATTTCCTCCTAAGTCTTCCCTCTGTGAATTTTTTCCCTATAATTTCA	TTTC	1967
Db	181	AGATAACTTTTTATTTCCTCCTAAGTCTTCCCTCTGTGAATTTTTTCCCTATAATTTCA	TTTC	240
OY	1968	TTGTTTTTTCATCTCATTTATCCAAGTTCTGCAGACCACACAGAAGTTCCTCATGGCTC		2027
Db	241	TTGTTTTTTCATCTCATTTATCCAAGTTCTGCAGACCACACAGAAGTTCCTCATGGCTC		300
OY	2028	TTTAGATGAATAGAAAGTTCAGGGTCCCCTCACCTCTAGTCACTAAAGAAGATTCTACC		2087
Db	301	TTTAGATGAATAGAAAGTTCAGGGTCCCCTCACCTCTAGTCACTAAAGAAGATTCTACTCT		360
OY	2088	CCCAGCCCCAGAAAAGGTGATTCCTCTCTTACCATTCTCGGGACCTTAGCTTAATTAGG		2147
Db	361	CCCAGCCCCAGAAAAGGTGATTCCT-TCCTTACCATTCTCGGGACCTTAGCTTAATTAGG		419
OY	2148	TACCTTATTTAACAGGAAGTGTAAAGGTACCTTCTCTGTGGAACAATCTGCAATGTCTAAA		2207
Db	420	TACCTTATTTAACAGGAAGTGTAAAGGTACCTTCTCTGTGGAACAATCTGCAATGTCTAAA		479
OY	2208	TGCGCCTTAAAGAGGCCCATTTCTTAGCTGCTGTAATCAGTGCTCTTCACTCTCTCAGAG		2267
Db	480	TGCGCCTTAAAGAGGCCCATTTCTTAGCTGCTGTAATCAGTGCTCTTCACTCTCTCAGAG		539
OY	2268	AAGCAGGGATGGTACCTTACCCGGCAGGTAGGTTAGATGTGGGTGGTCATGTTAATTTCC		2327
Db	540	AAGCAGGGATGGTACCTTACCCGGCAGGTAGGTTAGATGTGGGTGGTCATGTTAATTTCC		599
OY	2328	CTTAGAAGTTCAGGCCCTGTTCCTGCGTAAAGGTGATGTCCAGTTCAGAGATGTGT		2387
Db	600	CTTAGAAGTTCAGGCCCTGTTCCTGCGTAAAGGTGATGTCCAGTTCAGAGATGTGT		659
OY	2388	ATAATGAGCATGGCTGTGTTAAGATCAGGAGGCCCACTTGATTTTATAGTATAGCCCTTCC		2447
Db	660	ATAATGAGCATGGCTGTGTTAAGATCAGGAGGCCCACTTGATTTTATAGTATAGCCCTTCC		719
OY	2448	TCCACTCCCAACCAAGACTTGCTCATTTTTCGAGTTTAACTAGACTACACTCTATT		2503
Db	720	TCCACTCCCAACCAAGACTTGCTCATTTTTCGAGTTTAACTAGACTACACTCTATT		775

RESULT	11
AX069339	
LOCUS	AX069339
DEFINITION	Sequence 3 from Patent WO0102600.
ACCESSION	AX069339
VERSION	AX069339.1 GI:12579202
KEYWORDS	
SOURCE	human.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 2226)
AUTHORS	Yuan,C.S.
TITLE	Detection of analytes using attenuated enzymes
JOURNAL	Patent: WO 0102600-A 3 11-JAN-2001;
FEATURES	GENERAL ATOMICS (US) #3 Location/Qualifiers 1..2226 /organism="Homo sapiens" /db_xref="taxon:9606" 1..2226 /note="Polynucleotide encoding human S-adenosyl-5-homocysteine hydrolase (SAHH) derived from bladder; n=a, c, g, or t"
BASE COUNT	649 a 455 c 435 g 675 t 12 others
ORIGIN	
Query Match	21.6%; Score 553; DB 6; Length 2226;

Accession	Gene	Species	Chromosome	Position	Length	Orientation	Sequence
1728	GTTGCCAGCTTGCATCTGCCATCATTTGATGCCACCCTTACAGAGCTGACAGATGACCAA	Human	1	1728	1787	+	
1728	GTTGCCAGCTTGCATCTGCCATCATTTGATGCCACCCTTACAGAGCTGACAGATGACCAA	Human	1	1728	1787	+	
1788	GCAAAATATCTGGAGCTCAACAAAATGGGCCATTCAAACTAATTATTACAGATACTAA	Human	1	1788	1847	+	
1788	GCAAAATATCTGGAGCTCAACAAAATGGGCCATTCAAACTAATTATTACAGATACTAA	Human	1	1788	1847	+	
1848	TGGACCATCTACTACCAAGACAGCTCCACCTGAACACACACACTCTAAAGAAATATT	Human	1	1848	1907	+	
1848	TGGACCATCTACTACCAAGACAGCTCCACCTGAACACACACACTCTAAAGAAATATT	Human	1	1848	1907	+	
1908	AGATACTTTTATTTCTCTTCTTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	Human	1	1908	1967	+	
1908	AGATACTTTTATTTCTCTTCTTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	Human	1	1908	1967	+	
1968	TTGTTTTTTCATCTCATTTATCCAGTTCTGCAGACACACAGAACTTGCTCATGGCTC	Human	1	1968	2027	+	
1968	TTGTTTTTTCATCTCATTTATCCAGTTCTGCAGACACACAGAACTTGCTCATGGCTC	Human	1	1968	2027	+	
2028	TTTATGATGAATAGAAAGTTCAGGGTCCCTCAGCTCTAGTCACTAAAGAAAGATTTT	Human	1	2028	2087	+	
2028	TTTATGATGAATAGAAAGTTCAGGGTCCCTCAGCTCTAGTCACTAAAGAAAGATTTT	Human	1	2028	2087	+	
301	TTTATGATGAATAGAAAGTTCAGGGTCCCTCAGCTCTAGTCACTAAAGAAAGATTTT	Human	1	301	360	+	
301	TTTATGATGAATAGAAAGTTCAGGGTCCCTCAGCTCTAGTCACTAAAGAAAGATTTT	Human	1	301	360	+	
2088	CCAGCCAGAAAGGTGATCTCTCTCTTACCATTTCTGGGAGCTTTAGTCTTAATTAGG	Human	1	2088	2147	+	
2088	CCAGCCAGAAAGGTGATCTCTCTCTTACCATTTCTGGGAGCTTTAGTCTTAATTAGG	Human	1	2088	2147	+	
361	CCAGCCAGAAAGGTGATCTCTCTTACCATTTCTGGGAGCTTTAGTCTTAATTAGG	Human	1	361	419	+	
361	CCAGCCAGAAAGGTGATCTCTTACCATTTCTGGGAGCTTTAGTCTTAATTAGG	Human	1	361	419	+	
2148	TACCTTATTAACAGGAAATGCTAAGGTACCTTCTCTGTGGAACAATCTGCAATGCTAA	Human	1	2148	2207	+	
2148	TACCTTATTAACAGGAAATGCTAAGGTACCTTCTCTGTGGAACAATCTGCAATGCTAA	Human	1	2148	2207	+	
420	TACCTTATTAACAGGAAATGCTAAGGTACCTTCTCTGTGGAACAATCTGCAATGCTAA	Human	1	420	479	+	
420	TACCTTATTAACAGGAAATGCTAAGGTACCTTCTCTGTGGAACAATCTGCAATGCTAA	Human	1	420	479	+	
2208	TCGCTTAAAGAGAGCCCATTTCTTAGCTGCTGAATCAGTGTCTTCACTTCTTACAGAG	Human	1	2208	2267	+	
2208	TCGCTTAAAGAGAGCCCATTTCTTAGCTGCTGAATCAGTGTCTTCACTTCTTACAGAG	Human	1	2208	2267	+	
480	TCGCTTAAAGAGAGCCCATTTCTTAGCTGCTGAATCAGTGTCTTCACTTCTTACAGAG	Human	1	480	539	+	
480	TCGCTTAAAGAGAGCCCATTTCTTAGCTGCTGAATCAGTGTCTTCACTTCTTACAGAG	Human	1	480	539	+	
2268	AAGCAGGATGCTACCTACCCGCGAGTAGGTAGATGTGGTGTGCATGTTAATTGCC	Human	1	2268	2327	+	
2268	AAGCAGGATGCTACCTACCCGCGAGTAGGTAGATGTGGTGTGCATGTTAATTGCC	Human	1	2268	2327	+	
540	AAGCAGGATGCTACCTACCCGCGAGTAGGTAGATGTGGTGTGCATGTTAATTGCC	Human	1	540	599	+	
540	AAGCAGGATGCTACCTACCCGCGAGTAGGTAGATGTGGTGTGCATGTTAATTGCC	Human	1	540	599	+	
2328	CTTGAAGTCTCCAGGCCCTGTTCTCTGCGTAAAGGTGTATGTCAGTTCAGAGATGTGT	Human	1	2328	2387	+	
2328	CTTGAAGTCTCCAGGCCCTGTTCTCTGCGTAAAGGTGTATGTCAGTTCAGAGATGTGT	Human	1	2328	2387	+	
600	CTTGAAGTCTCCAGGCCCTGTTCTCTGCGTAAAGGTGTATGTCAGTTCAGAGATGTGT	Human	1	600	659	+	
600	CTTGAAGTCTCCAGGCCCTGTTCTCTGCGTAAAGGTGTATGTCAGTTCAGAGATGTGT	Human	1	600	659	+	
2388	ATAATGAGCATGGCTTGTAAAGATCAGGAGGCCCACTTGGAATTATAGTATAGCCCTTGC	Human	1	2388	2447	+	
2388	ATAATGAGCATGGCTTGTAAAGATCAGGAGGCCCACTTGGAATTATAGTATAGCCCTTGC	Human	1	2388	2447	+	
660	ATAATGAGCATGGCTTGTAAAGATCAGGAGGCCCACTTGGAATTATAGTATAGCCCTTGC	Human	1	660	719	+	
660	ATAATGAGCATGGCTTGTAAAGATCAGGAGGCCCACTTGGAATTATAGTATAGCCCTTGC	Human	1	660	719	+	
2448	TCGACTCCCAACAGACTTGCCTATTTTCGAGTTTAACTAGACTACACTCTATT	Human	1	2448	2503	+	
2448	TCGACTCCCAACAGACTTGCCTATTTTCGAGTTTAACTAGACTACACTCTATT	Human	1	2448	2503	+	
720	TCGACTCC						

COMMENT

humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Aug 19, 2002 this sequence version replaced g1:13396588.

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP database can be found at

http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr1
RP11-180N18 is from the library RPCI-11.1 constructed by the group of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm

VECTOR: pBACE3.6

Draft Sequence Produced by Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
http://www.seq.wi.mit.edu

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

FEATURES

source

Location/Qualifiers
1..81378
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-180N18"
/clone_1lb="RPCI-11.1"
/clone_1lb="RPCI-11.1"

BASE COUNT 21766 a 17337 c 18556 g 23719 t
ORIGIN

Query Match 17.4%; Score 446; DB 9; Length 81378;
Best Local Similarity 99.6%; Pred. No. 2.7e-248;

Matches 666; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

OY 1835 TTACAGATTAATGACCATCTACCAAGACAGTCCACCTGAACACACACTCTAA 1894
DB 49283 TTACAGATTAATGACCATCTACCAAGACAGTCCACCTGAACACACACTCTAA 49342
OY 1895 GAAATATTTTAAAGATACTTTATTTCTCTACTCTCTTCTGATTTTTC 1954
DB 49343 GAAATATTTTAAAGATACTTTATTTCTCTACTCTCTTCTGATTTTTC 49402
OY 1955 TATAATTCATCTCTGTTTTCATCTCATTAATCCAGTCTGCAGACACACAGAACT 2014
DB 49403 TATAATTCATCTCTGTTTTCATCTCATTAATCCAGTCTGCAGACACACAGAACT 49462
OY 2015 TGCTTCATGCTCTTTAGATGAATAGAGTTCAGGCTCCCTCACTAGTCACTAAGA 2074
DB 49463 TGCTTCATGCTCTTTAGATGAATAGAGTTCAGGCTCTCACTAGTCACTAAGA 49522
OY 2075 AGGATTTTACTCCCCAGCCAGAAAGTGATCTCTCTTACCATTTCTGGGACTTT 2134
DB 49523 AGGATTTTACTCTCCAGCCAGAAAGTGATCTT-TCCTTACCATTTCTGGGACTTT 49581
OY 2135 AGTCTTAATAGTACCTTAATTAACAGGAATGCTAAGTACCTTCTGTGGAACAATC 2194

DB 49582 AGCTTAATTAAGTACCTTAATTAACAGGAATGCTAAGTACCTTCTGTGGAACAATC 49641
OY 2195 TGCAATGCTTAATGCGCTTAAGAGCCCATTTCTAGCTGCTGAATCAGTCTCTT 2254
DB 49642 TGCAATGCTTAATGCGCTTAAGAGCCCATTTCTAGCTGCTGAATCAGTCTCTT 49701
OY 2255 CACTTCTCAGAGAAGCAGGATGTACTACCCGGCAGGTAGTTAGATGTGGTGTG 2314
DB 49702 CACTTCTCAGAGAAGCAGGATGTACTACCCGGCAGGTAGTTAGATGTGGTGTG 49761
OY 2315 CATGTTAATTTCCCTTAGAAGTCCAGCCCTGTTTCTGCGTAAGGTGTAATCCAG 2374
DB 49762 CATGTTAATTTCCCTTAGAAGTCCAGCCCTGTTTCTGCGTAAGGTGTAATCCAG 49821
OY 2375 TTCAGAGATGTGTAATATGAGCATGGCTTGTAAAGATCAGAGGCCACTTGTATTA 2434
DB 49822 TTCAGAGATGTGTAATATGAGCATGGCTTGTAAAGATCAGAGGCCACTTGTATTA 49881
OY 2435 GTATAGCCCTTCTCTCCACTCCACAGAGACTTGTCTATTTTTCAGATTCTAGACTA 2494
DB 49882 GTATAGCCCTTCTCTCCACTCCACAGAGACTTGTCTATTTTTCAGATTCTAGACTA 49941
OY 2495 CACTCTAAT 2503
DB 49942 CACTCTAAT 49950

RESULT 13

AL360171

ID AL360171 standard; DNA; HTG; 183584 BP.

XX AC AL360171; AC025865;

XX SV AL360171.14

XX DT 01-JUL-2000 (Rel. 64, Created)

DT 19-APR-2001 (Rel. 67, Last updated, Version 23)

XX DE Human DNA sequence *** SEQUENCING IN PROGRESS *** from clone RP11-180N18

XX KW HTG; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP; HTGS_PHASE2.

XX OS Homo sapiens (human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

XX RN [1]

RA Donnelly S.;

RT Submitted (07-APR-2001) to the EMBL/GenBank/DBJ databases.
RL Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

CC Draft Sequence Produced by Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

CC http://www.seq.wi.mit.edu

CC ----- Genome Center

CC Center: Sanger Centre

CC Center code: SC

CC Web site: http://www.sanger.ac.uk

CC Contact: humquery@sanger.ac.uk

CC ----- Project Information

CC Center project name: ba180N18

CC ----- Summary Statistics

CC Assembly program: XGAP4; version 4.5

CC Sequencing vector: M13; M77815; 25% of reads

CC Sequencing vector: plasmid; L08752; 74% of reads

CC Chemistry: Dye-terminator ABI; 0% of reads

CC Chemistry: Dye-terminator ET-amersham; 21% of reads

CC Chemistry: Dye-terminator Big Dye; 78% of reads

CC Consensus quality: 183396 bases at least Q40
CC Consensus quality: 183541 bases at least Q30
CC Consensus quality: 183560 bases at least Q20
CC Insert size: 183584; sum-of-ctrls
CC Insert size: 191911; 8.1% error; agarose-fp
CC Quality coverage: 10.84x in Q20 bases; sum-of-ctrls
CC Quality coverage: 10.92x in Q20 bases; agarose-fp
CC -----
CC * NOTE: This is a 'working draft' sequence. It currently
CC consists of 1 contigs. The true order of the pieces is
CC not known and their order in this sequence record is
CC arbitrary. Where the contigs adjacent to the vector can
CC be identified, they are labelled with 'clone-end' in the
CC feature table. Some order and orientation information
CC can tentatively be deduced from paired sequencing reads
CC which have been identified to span the gap between two
CC contigs. These are labelled as part of the same
CC 'fragment_chain', and the order and relative orientation
CC of the pieces within a fragment_chain is reflected in
CC this file. Gaps between the contigs are represented as
CC * runs of N, but the exact sizes of the gaps are unknown.
CC * This record will be updated with the finished sequence as
CC soon as it is available and the accession number will be
CC preserved.
CC 1 183584 contig of 183584 bp in length
XX
FH Key Location/Qualifiers
FH
FT source 1.183584
FT /db_xref="taxon:9606"
FT /organism="Homo sapiens"
FT /clone="RP11-180N18"
FT /clone_11b="RPC1-11.1"
FT misc_feature 1.183584
FT /note="assembly_fragment:00420"
XX
SQ Sequence 183584 BP; 45791 A; 43162 C; 44819 G; 49812 T; 0 other;
Query Match 17.4%; Score 446; DB 30; Length 183584;
Best Local Similarity 99.6%; Pred. No. 2.7e-248;
Matches 666; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1835 TTACAGATCTAATGACCACTACTACCAAGCAGCAGTCCACCTGAACACACACTCTAA 1894
DB 49283 TTACAGATCTAATGACCACTACTACCAAGCAGCAGTCCACCTGAACACACACTCTAA 49342
QY 1895 GAAATATTTTAAAGATTAATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1954
DB 49343 GAAATATTTTAAAGATTAATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 49402
QY 1955 TATAATTTTCATCTCTGTTTTCATCTCTATTCACCAAGTTCGACAGACACAGGAAGT 2014
DB 49403 TATAATTTTCATCTCTGTTTTCATCTCTATTCACCAAGTTCGACAGACACAGGAAGT 49462
QY 2015 TGCTTCATGGCTCTTTAGATGAATAAGATTGAGGTCCTCCTCCTCCTCCTCCTCCT 2074
DB 49463 TGCTTCATGGCTCTTTAGATGAATAAGATTGAGGTCCTCCTCCTCCTCCTCCTCCT 49522
QY 2075 AGGATTTTACTCCCCAGCCAGCAAGAGTGATCTCTCTCTTACCAATTTCTGGGACTTT 2134
DB 49523 AGGATTTTACTCCCCAGCCAGCAAGAGTGATCTCTCTTACCAATTTCTGGGACTTT 49581
QY 2135 AGCTTAATTAAGTACCTTAATTAACAGGAATGCTAAGGTACCTCTCTGTGGAACAATC 2194
DB 49582 AGCTTAATTAAGTACCTTAATTAACAGGAATGCTAAGGTACCTCTCTGTGGAACAATC 49641
QY 2195 TGCAATGCTTAATCGCCTTAAGAGCCCAATTTCTTAGCTGCTGAATTCAGTGTCTTT 2254
DB 49642 TGCAATGCTTAATCGCCTTAAGAGCCCAATTTCTTAGCTGCTGAATTCAGTGTCTTT 49701
QY 2255 CACTTCTCAGAGAGAGGAGGTGCTACCTACCCGCGAGGTAGGTAGATGTGGGTGTG 2314
|||||

DB 49702 CACTTCTCAGAGAGAGGAGGTGCTACCTACCCGCGAGGTAGGTAGATGTGGGTGTG 49761
QY 2315 CATGTTAATTTCCCTTAGAAGTCCAGGCCCTGTTCTCGGTAAGGTGTAATGCCAG 2374
DB 49762 CATGTTAATTTCCCTTAGAAGTCCAGGCCCTGTTCTCGGTAAGGTGTAATGCCAG 49821
QY 2375 TTCAGAGATGTGTATTAATGAGCATGGCTGTTAAGATCAGAGAGGCCACTGTGATTTATA 2434
DB 49822 TTCAGAGATGTGTATTAATGAGCATGGCTGTTAAGATCAGAGAGGCCACTGTGATTTATA 49881
QY 2435 GTATAGCCCTTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 2494
DB 49882 GTATAGCCCTTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 49941
QY 2495 CACTCTAAT 2503
DB 49942 CACTCTAAT 49950
RESULT 14
AX408384/c 376 bp DNA linear PAT 14-JUN-2002
LOCUS
DEFINITION Sequence 1031 from Patent WO0229103.
ACCESSION AX408384
VERSION AX408384.1 GI:21441089
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1
TITLE Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.
JOURNAL Gene expression profiles in liver cancer
Patent: WO 0229103-A 1031 11-APR-2002;
GENE LOGIC INC (US)
FEATURES
source 1.376
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. AA432168"
BASE COUNT 125 a 78 c 81 g 92 t
ORIGIN
Query Match 12.9%; Score 330; DB 6; Length 376;
Best Local Similarity 100.0%; Pred. No. 1.7e-180;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2174 TACCTTCTCTGTGGAACAATCTGCAATGCTTAATCGCCTTAAGAGAGGCCATTTCTTAG 2233
DB 376 TACCTTCTCTGTGGAACAATCTGCAATGCTTAATCGCCTTAAGAGAGGCCATTTCTTAG 317
QY 2234 CTGCTGAATCACTGCTCTCTTCACTTCTTCAAGAGAGAGGAGGTGCTACCTACCCGCGAG 2293
DB 316 CTGCTGAATCACTGCTCTCTTCACTTCTTCAAGAGAGAGGAGGTGCTACCTACCCGCGAG 2257
QY 2294 GTAGGTAGATGTGGGTGTGTCATGTTAATTTCCCTTAGAAGTCCAGGCCCTGTTTCTT 2353
DB 256 GTAGGTAGATGTGGGTGTGTCATGTTAATTTCCCTTAGAAGTCCAGGCCCTGTTTCTT 2317
QY 2354 GCGTAAGGTGATGTGTCAGTTCAGAGATGTGTAATGAGCATGGCTGTGTAAGATCA 2413
DB 196 GCGTAAGGTGATGTGTCAGTTCAGAGATGTGTAATGAGCATGGCTGTGTAAGATCA 2377
QY 2414 GGAGGCCCACTTGATTTATAGTATAGCCCTTCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 2473
DB 136 GGAGGCCCACTTGATTTATAGTATAGCCCTTCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 2437
QY 2474 TTGAGTTTCTTACTAGACTACACTTAAT 2503
DB 76 TTGAGTTTCTTACTAGACTACACTTAAT 47
RESULT 15

AX366273/c
LOCUS AX366273 292 bp DNA linear PAT 15-FEB-2002
DEFINITION Sequence 40 from Patent WO0206317.
ACCESSION AX366273
VERSION AX366273.1 GI:18697699
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1
Mitcham,J.L., King,G.E., Algate,P.A., Fling,S.P., Retter,M.W.,
Fanger,G.R., Reed,S.G., Vedvick,T.S., Carter,D., Hill,P. and
Albone,E.
TITLE Compositions and methods for the therapy and diagnosis of ovarian
cancer
JOURNAL Patent: WO 0206317-A 40 24-JAN-2002;
CORIXA CORPORATION (US)
FEATURES
Source 1..292
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 98 a 68 c 60 g 66 t
ORIGIN
Query Match 9.2%; Score 237; DB 6; Length 292;
Best Local Similarity 100.0%; Pred. No. 4e-126;
Matches 237; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2267 GAAGCAGGGATGTACCTACCCGGCAGGTAGGTAGATGTGGGTGTCATGTTAATTTC 2326
DB 276 GAAGCAGGGATGTACCTACCCGGCAGGTAGGTAGATGTGGGTGTCATGTTAATTTC 217
OY 2327 CCTTAGAGTTCACAGCCCTGTTTCTGCGTAAAGGTGTATGTCACAGTTCAGAGATGTG 2386
DB 216 CCTTAGAGTTCACAGCCCTGTTTCTGCGTAAAGGTGTATGTCACAGTTCAGAGATGTG 157
OY 2387 TATTAATGAGCATGGCTGTTAAGATCAGAGGCCCACTTGATTATATAGTACCCCTTC 2446
DB 156 TATTAATGAGCATGGCTGTTAAGATCAGAGGCCCACTTGATTATATAGTACCCCTTC 97
OY 2447 CTCACATCCACACAGACTGTCTCATTTTCGAGTTTAACTAGACTACACTCTATT 2503
DB 96 CTCACATCCACACAGACTGTCTCATTTTCGAGTTTAACTAGACTACACTCTATT 40

Search completed: April 20, 2003, 21:15:35
Job time : 6890 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2003, 17:23:45 ; Search time 543 Seconds
(without alignments)
10629.598 Million cell updates/sec

Title: US-09-782-051-1
Perfect score: 2563
Sequence: 1 ggcgcggcgcagctcgcagct.....aaaaaaaaaaaaaaaaaaaaa 2563

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_101002:*

1:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2563	100.0	2563	19 AAV28617	Nucleotide sequenc
2	2178	85.0	3634	23 ABV23195	Human prostate exp
3	2178	85.0	3634	23 ABV29032	Human prostate exp
4	1522	59.4	3616	20 AAV73924	Human SAHH DNA #1.
5	570	22.2	721	21 AAF16004	Human prostate can
6	553	21.6	2226	20 AAV73925	Human SAHH DNA #2.
7	422	16.5	504	23 ABV44141	Human prostate exp
8	370	14.4	636	21 AAZ80766	Human COL1 cancer
9	330	12.9	376	24 ABN94533	Gene #1031 used to

C	10	320	12.5	349	21	AAA44806	Human secreted exp
C	11	237	9.2	292	21	AAA69730	Human ovarian carc
C	12	237	9.2	292	24	ABN72624	Ovarian carcinoma
C	13	224	8.7	553	24	ABL79045	Human ovarian carc
C	14	190	7.4	1877	18	AAV06354	AL-2-long (AL-21)
	15	168	6.6	384	24	ABN93848	Gene #346 used to
	16	118	4.6	312	23	ABV35309	Human prostate exp
	17	118	4.6	435	23	ABV14215	Human prostate exp
	18	117	4.6	326	21	AAC28464	Human secreted pro
	19	95	3.7	278	24	ABN73754	Bovine embryonic g
	20	82	3.2	328	23	ABV05046	Human prostate exp
	21	69	2.7	433	21	AAC21046	Human secreted pro
	22	66	2.6	138	16	AAT25056	Human gene signatu
	23	60	2.3	60	24	ABN33274	Human spliced tran
C	24	26	1.0	207	22	AAH71190	Human cervical can
	25	26	1.0	259	18	AAT91303	Human H1075-1 secr
	26	26	1.0	259	18	AAV00423	3' fragment of clo
C	27	26	1.0	315	23	ABV35453	Human prostate exp
C	28	26	1.0	352	23	ABV14369	Human prostate exp
C	29	26	1.0	451	22	AAI09877	Human breast cance
C	30	26	1.0	451	22	AAI18549	Human breast cance
C	31	26	1.0	452	22	AAI18548	Human breast cance
C	32	26	1.0	492	22	AAI09876	Human breast cance
C	33	26	1.0	863	18	AAT91302	Human H1075-1 secr.
C	34	26	1.0	987	24	ABN98493	Arabidopsis thaila
C	35	26	1.0	1165	24	AAD23915	Human dual-specific
	36	26	1.0	178896	24	ABQ88146	Human osteoblast d
C	37	25	1.0	90	23	ABV35515	Human prostate exp
	38	25	1.0	187	24	ABK39436	DNA encoding lung
	39	25	1.0	232	23	ABV19107	Human prostate exp
	40	25	1.0	241	23	ABV08717	Human prostate exp
C	41	25	1.0	294	24	ABL87534	Human ovarian can
C	42	25	1.0	353	21	AAC77649	Human cancer assoc
C	43	25	1.0	373	22	AAH68727	Human cervical can
C	44	25	1.0	396	22	AAK69700	Human immune/haema
	45	25	1.0	410	22	AAI91275	Human polynucleoti

ALIGNMENTS

RESULT 1	AAV28617	standard; cDNA; 2563 BP.
ID	AAV28617;	
AC	AAV28617;	
XX		
XX	21-AUG-1998	(first entry)
DT		
XX		
DE	Nucleotide sequence of the ODD4b5.3 enzyme.	
XX		
KW	DD4b9.3 enzyme; S-adenosyl-L-homocysteine hydrolase-type activity;	
KW	ANCY-type activity; stimulation; inhibition; dendritic cell; cancer;	
KW	autoimmune disease; transplantation; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	3..1847
FT		/*tag- a
FT		/product- "DD4b5.3 enzyme"
XX		
PN	W09814562-A1.	
XX		
PD	09-APR-1998.	
XX		
PF	06-OCT-1997;	97WO-NZ00133.
XX		
PR	04-OCT-1996;	96NZ-0299507.
XX		
PA	(HART/) HART D.N J.	
XX		
PI	Hart DNT;	

XX WPI: 1998-240074/21.
 DR P-PSDB; AAW56097.
 XX New isolated S-adenosyl-L-homocysteine hydrolase enzyme - is used to
 PT develop products which can be used in the treatment of e.g.
 PT auto-immune disease, transplantations or cancers
 XX
 PS Claim 7; Fig 1; 33pp; English.
 XX
 CC This is the nucleotide sequence of the DD4b9.3 enzyme which has
 CC S-adenosyl-L-homocysteine hydrolase (AHCY)-type activity. Its
 CC products can be used to identify substances which have a stimulatory
 CC or inhibitory effect on the enzyme activity. Such substances can be
 CC used to modulate dendritic cell (DC) function and for immunomodulation.
 CC They can be used in the treatment of e.g. autoimmune diseases,
 CC transplantations or cancers. The products can also be used for
 CC detection and diagnosis.
 CC
 XX
 SO Sequence 2563 BP; 646 A; 604 C; 677 G; 636 T; 0 other;

Query Match 100.0%; Score 2563; DB 19; Length 2563;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGCGGGCAGGTCGGAGCTCGAGCTGCTGCTTCTGTTCTGTGGCCACCGTGGCT 60
 DB 1 GGGCGGGCAGGTCGGAGCTCGAGCTGCTGCTTCTGTTCTGTGGCCACCGTGGCT 60
 QY 61 GTCCGCTGCTTGGGCTGCGAACAAGAGCGCTGGCCACAGCACCCTCAGAACCGA 120
 DB 61 GTCCGCTGCTTGGGCTGCGAACAAGAGCGCTGGCCACAGCACCCTCAGAACCGA 120
 QY 121 CGCAGCTCGACGAGGGGGCGCAGAGGGTGGGGCGATCGCGTGTGAGAGGGCGCGCC 180
 DB 121 CGCAGCTCGACGAGGGGGCGCAGAGGGTGGGGCGATCGCGTGTGAGAGGGCGCGCC 180
 QY 181 GGGCAGGG 240
 DB 181 GGGCAGGG 240
 QY 241 GGGCGGG 300
 DB 241 GGGCGGG 300
 QY 301 TGAAGCAGGCCAAGAGATCGAGAGCGCCGAGAAGTACTCTTCATGACCACCTCACA 360
 DB 301 TGAAGCAGGCCAAGAGATCGAGAGCGCCGAGAAGTACTCTTCATGACCACCTCACA 360
 QY 361 AGGGCCCAAGAGCAATCCAGTTGCTGATGACATGACAGAGTTCACCAATTCCTCCA 420
 DB 361 AGGGCCCAAGAGCAATCCAGTTGCTGATGACATGACAGAGTTCACCAATTCCTCCA 420
 QY 421 CCAAACTGGCCGAAGATCTTGTCTGCTGATCTCACAGTCTCCACTGACAGCTACA 480
 DB 421 CCAAACTGGCCGAAGATCTTGTCTGCTGATCTCACAGTCTCCACTGACAGCTACA 480
 QY 481 GTTCAGCTGCTCTACACAGATAGCTGATGATGAGGTTTCTCCCGAGAGAGCAGC 540
 DB 481 GTTCAGCTGCTCTACACAGATAGCTGATGATGAGGTTTCTCCCGAGAGAGCAGC 540
 QY 541 AAAAGCACTCCAAGGGCAGCAGCAATTTCTGCTGAAGAATCAAGCAGGCAATTTG 600
 DB 541 AAAAGCACTCCAAGGGCAGCAGCAATTTCTGCTGAAGAATCAAGCAGGCAATTTG 600
 QY 601 GACGCCGGAGATGAGATGACAGAGCAAGATGCTGCTCTGATTTCACTCAGAGAAC 660
 DB 601 GACGCCGGAGATGAGATGACAGAGCAAGATGCTGCTCTGATTTCACTCAGAGAAC 660
 QY 661 GTGCTCAGGGGGAGAGCCCTTGGCTGCTGCTAAATAAGTGGGCTGTACACACATCAG 720
 DB 661 GTGCTCAGGGGGAGAGCCCTTGGCTGCTGCTAAATAAGTGGGCTGTACACACATCAG 720

QY 721 CCCAGACAGCGGTGTGATGAGACACTCTGTGCCCCCTGGGGGCTCAGTGGCCGTCTG 780
 DB 721 CCCAGACAGCGGTGTGATGAGACACTCTGTGCCCCCTGGGGGCTCAGTGGCCGTCTG 780
 QY 781 CTTGTAACATCTACTCAACTCAGAAATGAGTAGCTGCAGCAGCTGGCTGAGGCTGAGT 840
 DB 781 CTTGTAACATCTACTCAACTCAGAAATGAGTAGCTGCAGCAGCTGGCTGAGGCTGAGT 840
 QY 841 CAGTGTGCTTGGAGGGCGAGTCAGAGATGACTTCTGTGTGTATGACCGCTGTG 900
 DB 841 CAGTGTGCTTGGAGGGCGAGTCAGAGATGACTTCTGTGTGTATGACCGCTGTG 900
 QY 901 TGAACATGATGGGTGGCAGGCCAATGATCTGATGATGGGGAGAGACTTAACCACT 960
 DB 901 TGAACATGATGGGTGGCAGGCCAATGATCTGATGATGGGGAGAGACTTAACCACT 960
 QY 961 GGGTTATAGAAGATATCCAAACGCTTTAAGAGATCCGAGGATTTGGAAGAGAGCG 1020
 DB 961 GGGTTATAGAAGATATCCAAACGCTTTAAGAGATCCGAGGATTTGGAAGAGAGCG 1020
 QY 1021 TGACTGTGTTACAGAGCGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTCCGCCA 1080
 DB 1021 TGACTGTGTTACAGAGCGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTCCGCCA 1080
 QY 1081 TGAACGTCATGATTTCTGTTACCAACAGAGTTGATACTGTACTGCTGCCAGAAAT 1140
 DB 1081 TGAACGTCATGATTTCTGTTACCAACAGAGTTGATACTGTACTGCTGCCAGAAAT 1140
 QY 1141 CCATTTTGAATGGCTTGAAGAGACACAGATGTGTTGGTGGGAACAAGTGTG 1200
 DB 1141 CCATTTTGAATGGCTTGAAGAGACACAGATGTGTTGGTGGGAACAAGTGTG 1200
 QY 1201 TGTGTGCTATGTTGAGGTAGGCAAGGGCTGCTGTCTCTCAAGCTCTTGGAGCAA 1260
 DB 1201 TGTGTGCTATGTTGAGGTAGGCAAGGGCTGCTGTCTCTCAAGCTCTTGGAGCAA 1260
 QY 1261 TTGTCTACATTAACGAATCGAACCCCATGTGTCTGTCAGGCTGATGATGGTTCA 1320
 DB 1261 TTGTCTACATTAACGAATCGAACCCCATGTGTCTGTCAGGCTGATGATGGTTCA 1320
 QY 1321 GGGTGTAAAGCTAAATGAAGTCAATCCGCAAGTCGATGCTAATAACTTGACAGGAA 1380
 DB 1321 GGGTGTAAAGCTAAATGAAGTCAATCCGCAAGTCGATGCTAATAACTTGACAGGAA 1380
 QY 1381 ATAAGATGTAGTACACAGGGAGCAGCTTGATCGCATGAATAAGTGTATGATGCA 1440
 DB 1381 ATAAGATGTAGTACACAGGGAGCAGCTTGATCGCATGAATAAGTGTATGATGCA 1440
 QY 1441 ATATGGGCGCACTCCACACAGAAATCGATGTGACCAAGCTCCGCACTCCGAGCTGACGT 1500
 DB 1441 ATATGGGCGCACTCCACACAGAAATCGATGTGACCAAGCTCCGCACTCCGAGCTGACGT 1500
 QY 1501 GGGAGCAGTACGTTCTCAGGTGACCATGTCACTGCGCAGATGGCAAGGAGTTGCC 1560
 DB 1501 GGGAGCAGTACGTTCTCAGGTGACCATGTCACTGCGCAGATGGCAAGGAGTTGCC 1560
 QY 1561 TCCTGGCAGAGGCTGCTACTCAATTTGAGCTGCTCCACAGTTCCACCTTTGTTCTGT 1620
 DB 1561 TCCTGGCAGAGGCTGCTACTCAATTTGAGCTGCTCCACAGTTCCACCTTTGTTCTGT 1620
 QY 1621 CCATCAGAGCCACACACAGGCTTTGGCAGCTGATAGAACTCTATAATGACCCGAGGGG 1680
 DB 1621 CCATCAGAGCCACACACAGGCTTTGGCAGCTGATAGAACTCTATAATGACCCGAGGGG 1680
 QY 1681 GATACAGCAGAGTGTACTTCTTCTTAAGAAATGATGAATACGTGCGACGCTG 1740
 DB 1681 GATACAGCAGAGTGTACTTCTTCTTAAGAAATGATGAATACGTGCGACGCTG 1740
 QY 1741 ATCTGCAATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1800
 DB 1741 ATCTGCAATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1800
 QY 1801 GACTCAACAAAATGGCCATTCAAACTTAATTAAGATTAATGACCATATGACCATATCTAC 1860

DB 1801 GACTCACAATAATGGCCATTCACACCTAATATATACAGATACTAATGACCATACTAC 1860
QY 1861 CAAGAGCAGTCCACCTGAACACACACACTCTAAGAAATATTTTAAAGATACTTTAT 1920
DB 1861 CAAGAGCAGTCCACCTGAACACACACACTCTAAGAAATATTTTAAAGATACTTTAT 1920
QY 1921 TTTCTTCTTACTCTTCTCTCTCTGATTTTTCCTATATTTCACTTCTGTTTTCATC 1980
DB 1921 TTTCTTCTTACTCTTCTCTCTGATTTTTCCTATATTTCACTTCTGTTTTCATC 1980
QY 1981 TCATTATCCAGTCTGACAGACACACAGAACTGCTTCATGGCTCTTAGATGAATA 2040
DB 1981 TCATTATCCAGTCTGACAGACACACAGAACTGCTTCATGGCTCTTAGATGAATA 2040
QY 2041 GAAGTTCAGGTCCTCCTCACTCTAGTCACTAAGAAGATTTACTCCCCAGCCAGAAA 2100
DB 2041 GAAGTTCAGGTCCTCCTCACTCTAGTCACTAAGAAGATTTACTCCCCAGCCAGAAA 2100
QY 2101 GGTGATCTCTCTTACCATTTCTGGGACTTTAGTCTTAATAGTACCTTATTAACA 2160
DB 2101 GGTGATCTCTCTTACCATTTCTGGGACTTTAGTCTTAATAGTACCTTATTAACA 2160
QY 2161 GGAATGCTAAGTACCTCTCTGTGGAACAATCTGCAATGTCTAATCGCCTTAAAGA 2220
DB 2161 GGAATGCTAAGTACCTCTCTGTGGAACAATCTGCAATGTCTAATCGCCTTAAAGA 2220
QY 2221 GCCCATTTCTTAGCTGTGAATATCAGTCTCTTCACTTCTTCAAGAGCAGGGATGT 2280
DB 2221 GCCCATTTCTTAGCTGTGAATATCAGTCTCTTCACTTCTTCAAGAGCAGGGATGT 2280
QY 2281 ACCTACCCGGCAGGTAGGTAGATGTGGGTGTCATGTTAATTCCTTAGAAGTCCA 2340
DB 2281 ACCTACCCGGCAGGTAGGTAGATGTGGGTGTCATGTTAATTCCTTAGAAGTCCA 2340
QY 2341 AGCCCTGTTCTGCGTAAAGTGTATGTCCAGTTCAGAGATGTATTAATGAGCATGG 2400
DB 2341 AGCCCTGTTCTGCGTAAAGTGTATGTCCAGTTCAGAGATGTATTAATGAGCATGG 2400
QY 2401 CTTGTTAAGATCAGAGAGCCCACTTGATTTATAGTATAGCCCTTCCCACTCCCAACA 2460
DB 2401 CTTGTTAAGATCAGAGAGCCCACTTGATTTATAGTATAGCCCTTCCCACTCCCAACA 2460
QY 2461 GACTTGCATTTTTCGAGTTTAACTAGACTACACTCTATTTGAGTTAATTTGTCC 2520
DB 2461 GACTTGCATTTTTCGAGTTTAACTAGACTACACTCTATTTGAGTTAATTTGTCC 2520
QY 2521 TCTAGGATTTATTTCTGTGTCCAAAAA 2563
DB 2521 TCTAGGATTTATTTCTGTGTCCAAAAA 2563
RESULT 2
ABV23195
ID ABV23195 standard; cDNA; 3634 BP.
XX ABV23195;
AC
XX 16-SEP-2002 (first entry)
DT
XX Human prostate expression marker cDNA 23186.
DE
XX Human prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX Homo sapiens.
OS
XX WO200160860-A2.
PN
XX 23-AUG-2001.
PD
XX 20-FEB-2001; 2001WO-US05171.
PF
XX

PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer
PS Claim 1; Page 4168-4169; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 3634 BP; 876 A; 966 C; 869 G; 918 T; 5 other;
Query Match 85.0%; Score 2178; DB 23; Length 3634;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2498; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 1 GGGCGGGGAGGTCGAGCTCGAGCTGCTCTGTTCTCTTGTGGCCACCGTGGCT 60
DB 71 GGGCGGGGAGGTCGAGCTCGAGCTGCTCTGTTCTCTTGTGGCCACCGTGGCT 130
QY 61 GTCCGGCTGCTGGGCTGCCGACAGACAGAGCGGTGGCCACAGCAGCCTCAGAGCCGA 120
DB 131 GTCCGGCTGCTGGGCTGCCGACAGACAGAGCGGTGGCCACAGCAGCCTCAGAGCCGA 190
QY 121 CGCAGCTCGACGAGGGGCGGCGAGAGGCTGGCGATCGCGTGGAGGGCGCGCGC 180
DB 191 CGCAGCTCGACGAGGGGCGGCGAGAGGCTGGCGATCGCGTGGAGGGCGCGCGC 250
QY 181 GGGCAGGGGCGGCGGCGGAGAGGGGAGAGGGGCGGCGGCTCAGCCGCTGGCC 240
DB 251 GGGCAGGGGCGGCGGCGGAGAGGGGAGAGGGGCGGCGGCTCAGCCGCTGGCC 310
QY 241 GGGCGGGGGAATGTCGATGCTGACGCGATGCCGCTGCCCGGGTGGGAGAGAGC 300
DB 311 GGGCGGGGGAATGTCGATGCTGACGCGATGCCGCTGCCCGGGTGGGAGAGAGC 370
QY 301 TGAAGCAGGCGAAGAGATGAGGAGCGCGAGAGTACTCTTCATGGCCACCGTACCA 360
DB 371 TGAAGCAGGCGAAGAGATGAGGAGCGCGAGAGTACTCTTCATGGCCACCGTACCA 430
QY 361 AGGCGCCCAAGCAAAATCCAGTTTGTGATGATGACAGGAGTTTCAACCAATTCCCA 420
DB 431 AGGCGCCCAAGCAAAATCCAGTTTGTGATGATGACAGGAGTTTCAACCAATTCCCA 490
QY 421 CCAAACTGGCGAAGATCTTGTCTGCTGCTGATCTACAGTCTCCACTGACAGCTACA 480
DB 491 CCAAACTGGCGAAGATCTTGTCTGCTGCTGATCTACAGTCTCCACTGACAGCTACA 550

OY 481 GTTCAGCTGCATCCTACACAGATAGCTCTGATGATGAGTTTCTCCCGAGAGAAGCAGC 540
|||||
Db 551 GTTCAGCTGCATCCTACACAGATAGCTCTGATGATGAGTTTCTCCCGAGAGAAGCAGC 610
OY 541 AAACCAACTCCAAAGGCGACAGCAATTTCTGTGTGAAGAACATCAAGCAGCAGAAATTG 600
|||||
Db 611 AAACCAACTCCAAAGGCGACAGCAATTTCTGTGTGAAGAACATCAAGCAGCAGAAATTG 670
OY 601 GACGCCGGGAGATTGAGATTGCAGAGCAAGACATGTCTGCTCTGATTTCATCAGGAAC 660
|||||
Db 671 GACGCCGGGAGATTGAGATTGCAGAGCAAGACATGTCTGCTCTGATTTCATCAGGAAC 730
OY 661 GTGCTCAGGGGAGAGAGCCCTTGCTGCTGCTAAATAGTGGGCTGTACACATCAG 720
|||||
Db 731 GTGCTCAGGGGAGAGAGCCCTTGCTGCTGCTAAATAGTGGGCTGTACACATCAG 790
OY 721 CCCAGACAGCGGTGTTGATTGAGACACTCTGTGCCCTGGGGCTCAGTGGCGCTGTCTG 780
|||||
Db 791 CCCAGACAGCGGTGTTGATTGAGACACTCTGTGCCCTGGGGCTCAGTGGCGCTGTCTG 850
OY 781 CTGTGAACATCTACTCAACTCAGATGAAGTAGCTGCAGCACTGGCTGAGGCTGAGTTG 840
|||||
Db 851 CTGTGAACATCTACTCAACTCAGATGAAGTAGCTGCAGCACTGGCTGAGGCTGAGTTG 910
OY 841 CAGTGTTCGCTTGGAAGGGCGAGTCAGAGATGACTTCTGTGCTGTATGACCGCTGTG 900
|||||
Db 911 CAGTGTTCGCTTGGAAGGGCGAGTCAGAGATGACTTCTGTGCTGTATGACCGCTGTG 970
OY 901 TGAACATGATGGGTGGCAGGGCCAAATGATCTGATGATGGGGAGACCTTAACCCACT 960
|||||
Db 971 TGAACATGATGGGTGGCAGGGCCAAATGATCTGATGATGGGGAGACCTTAACCCACT 1030
OY 961 GGGTTTATAAGAGTATCCAAACGCTTTTAAGAAGATCCGAGCATTTGTGAAGAGCG 1020
|||||
Db 1031 GGGTTTATAAGAGTATCCAAACGCTTTTAAGAAGATCCGAGCATTTGTGAAGAGCG 1090
OY 1021 TGACTGCTGTTACAGAGCTGATCAAGCTCTCCAAGCTGGGAAGCTGTGTCCGGCCA 1080
|||||
Db 1091 TGACTGCTGTTACAGAGCTGATCAAGCTCTCCAAGCTGGGAAGCTGTGTCCGGCCA 1150
OY 1081 TGAACGTCAATGATTTCTGTACCAACAGAAAGTTGATTAATCTGTACTGCTGCCGAGAAT 1140
|||||
Db 1151 TGAACGTCAATGATTTCTGTGTACCAACAGAAAGTTGATTAATCTGTACTGCTGCCGAGAAT 1210
OY 1141 CCATTTTGGATGGCCTGAAGAGGACACAGATGTGATGTTGGTGGAAACAAGTGTGG 1200
|||||
Db 1211 CCATTTTGGATGGCCTGAAGAGGACACAGATGTGATGTTGGTGGAAACAAGTGTGG 1270
OY 1201 TGTGTGCTATGTTGAGTAGGCAAGGCTGCTGTGCTCTCAAGCTCTTGAGCAA 1260
|||||
Db 1271 TGTGTGCTATGTTGAGTAGGCAAGGCTGCTGTGCTCTCAAGCTCTTGAGCAA 1330
OY 1261 TTGCTTACATTACCGAAATGAGCCCATCTGTGCTCTGAGGCTGATGATGGTTCA 1320
|||||
Db 1331 TTGCTTACATTACCGAAATGAGCCCATCTGTGCTCTGAGGCTGATGATGGTTCA 1390
OY 1321 GGGTGGTAAAGCTAATGAAGTCAATCCGGCAAGTGTGATGATTAATCTTGACAGGAA 1380
|||||
Db 1391 GGGTGGTAAAGCTAATGAAGTCAATCCGGCAAGTGTGATGATTAATCTTGACAGGAA 1450
OY 1381 ATAAGAATGTAGTACACGGGAGCACTTGATCGCATGAAAAACAGTTGTATCGTATGCA 1440
|||||
Db 1451 ATAAGAATGTAGTACACGGGAGCACTTGATCGCATGAAAAACAGTTGTATCGTATGCA 1510
OY 1441 ATATGGGCCACTCCACACAGAAATCGATGTGACAGCCCTCGCACTCCGGAGCTGACGT 1500
|||||
Db 1511 ATATGGGCCACTCCACACAGAAATCGATGTGACAGCCCTCGCACTCCGGAGCTGACGT 1570
OY 1501 GGGAGCGAGTACGTTCTCAGGTGAGCATGTCATCTGGCAGAGTGGCAACGAGTTGTCC 1560
|||||
Db 1571 GGGAGCGAGTACGTTCTCAGGTGAGCATGTCATCTGGCAGAGTGGCAACGAGTTGTCC 1630
OY 1561 TCCTGGCAGAGGCTGCTACTCAATTGAGCTGCTCCACAGTTCCACCTTTGTCTGT 1620

Db 1631 TCCTGGCAGAGGCTGCTACTCAATTTGAGCTGCTCCACAGTTCCACCTTTGTCTGT 1690
OY 1621 CCATCAGAGCCACAACACAGGCTTTGGCACTGATGAACCTATATATGACCCGAGGGGC 1680
|||||
Db 1691 CCATCAGAGCCACAACACAGGCTTTGGCACTGATGAACCTATATATGACCCGAGGGGC 1750
OY 1681 GATACAGCAGGATGTGACTTGTCTTCCTAAGAAAAATGATGAATACGTTGCCAGCTTGC 1740
|||||
Db 1751 GATACAGCAGGATGTGACTTGTCTTCCTAAGAAAAATGATGAATACGTTGCCAGCTTGC 1810
OY 1741 ATCTGCCATCATTTGATGCCCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1800
|||||
Db 1811 ATCTGCCATCATTTGATGCCCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1870
OY 1801 GACTCAACAAAAATGGGCCATTCAAACCTAATTAATACAGATACTAATGACCAATCTAC 1860
|||||
Db 1871 GACTCAACAAAAATGGGCCATTCAAACCTAATTAATACAGATACTAATGACCAATCTAC 1930
OY 1861 CAAGACAGTCCACCTGAACCAACACACACTCTAAGAAAAATTTTAAAGATACTTTAT 1920
|||||
Db 1931 CAAGACAGTCCACCTGAACCAACACACACTCTAAGAAAAATTTTAAAGATACTTTAT 1990
OY 1921 TTCTCTTACTCCTTCTCTCTGATTTTTCCTAATAATTTCAATCTGTGTTTTCATC 1980
|||||
Db 1991 TTCTCTTACTCCTTCTCTCTGATTTTTCCTAATAATTTCAATCTGTGTTTTCATC 2050
OY 1981 TCATTAATCAAGTCTGACAGACACACAGAACTTCTCATGCTCTTAGATGAATA 2040
|||||
Db 2051 TCATTAATCAAGTCTGACAGACACACAGAACTTCTCATGCTCTTAGATGAATA 2110
OY 2041 GAAGTTCAGGGTCCCTCACTCTAGTCACTAAGAAAGATTTTACTCCCCAGCCAGAAA 2100
|||||
Db 2111 GAAGTTCAGGGTCCCTCACTCTAGTCACTAAGAAAGATTTTACTCTCCAGCCAGAAA 2170
OY 2101 GGTGATCTTCTCTTACCAATTTCTGGGACCTTAGTCTTAATAGGTACCTAATAACA 2160
|||||
Db 2171 GGTGATCTTCTCTTACCAATTTCTGGGACCTTAGTCTTAATAGGTACCTAATAACA 2229
OY 2161 GGAATGCTAAGGTACCTTCTCTGTGAACAACTGCAATGTCTAATGCTTAAGA 2220
|||||
Db 2230 GGAATGCTAAGGTACCTTCTCTGTGAACAACTGCAATGTCTAATGCTTAAGA 2289
OY 2221 GCCCATTTCTTAGCTGCTGAATCAGTGTCTTTCACTTCTCAGAGAGCAGGATGGT 2280
|||||
Db 2290 GCCCATTTCTTAGCTGCTGAATCAGTGTCTTTCACTTCTCAGAGAGCAGGATGGT 2349
OY 2281 ACCTACCCGCGAGGTAGTGTAGATGTGGGTGTCATGTTAATTTCCCTTAGAAGTTCCA 2340
|||||
Db 2350 ACCTACCCGCGAGGTAGTGTAGATGTGGGTGTCATGTTAATTTCCCTTAGAAGTTCCA 2409
OY 2341 AGCCCTGTTTCTGCGTAAAGGTGATGTGTCAGTTCAGAGATGTGTATATGAGCATGG 2400
|||||
Db 2410 AGCCCTGTTTCTGCGTAAAGGTGATGTGTCAGTTCAGAGATGTGTATATGAGCATGG 2469
OY 2401 CTGTTAAGATCAGAGGCGCCACTTGATTTATAGTATAGCCCTTCCCTCACCTCCACCA 2460
|||||
Db 2470 CTGTTAAGATCAGAGGCGCCACTTGATTTATAGTATAGCCCTTCCCTCACCTCCACCA 2529
OY 2461 GACTGCTCATTTTTCGAGTTTAACTAGACTACACTCTATT 2503
|||||
Db 2530 GACTGCTCATTTTTCGAGTTTAACTAGACTACACTCTATT 2572

RESULT 3
ABV29032
ID ABV29032 standard, cDNA; 3634 BP.
XX
AC ABV29032;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 29023.

Db 1451 ATAGAATGTAGTACACGGAGCACTGGATCGCATGAAAAACAGTTGTATCGTATGCA 1510
QY 1441 ATATGGCCACTCCACACAGAAATCGATGTGACCAGCCTCCGACTCCGGAGCTGACGT 1500
Db 1511 ATATGGCCACTCCACACAGAAATCGATGTGACCAGCCTCCGACTCCGGAGCTGACGT 1570
QY 1501 GGGAGCGAGTACGTTCTCAGGTGACCATGTGCTGCGCAGATGGCAACGAGTTGTCC 1560
Db 1571 GGGAGCGAGTACGTTCTCAGGTGACCATGTGCTGCGCAGATGGCAACGAGTTGTCC 1630
QY 1561 TCCTGGCAGAGGCTCGTCTACTCAATTTGAGCTGCTCCACAGTCCACCTTTGTCTGT 1620
Db 1631 TCCTGGCAGAGGCTCGTCTACTCAATTTGAGCTGCTCCACAGTCCACCTTTGTCTGT 1690
QY 1621 CCATCACAGCCACACACAGGCTTTGGCACTGATGAACCTCTATATGCAACCCGAGGGGC 1680
Db 1691 CCATCACAGCCACACACAGGCTTTGGCACTGATGAACCTCTATATGCAACCCGAGGGGC 1750
QY 1681 GATACACAGAGATGTGTACTGCTTCTTAAGAAAATGGATGAATACGTTGCCAGCTTGC 1740
Db 1751 GATACACAGAGATGTGTACTGCTTCTTAAGAAAATGGATGAATACGTTGCCAGCTTGC 1810
QY 1741 ATCTGCATCATTTGATGCCCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1800
Db 1811 ATCTGCATCATTTGATGCCCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1870
QY 1801 GACTCAACAAAATGGGCCATCAACCTAATATTATTACAGATACATAAGACCATACTAC 1860
Db 1871 GACTCAACAAAATGGGCCATCAACCTAATATTATTACAGATACATAAGACCATACTAC 1930
QY 1861 CAAGGACCACTCCACCTGAACCAACACACTCTAAAGAAATATTTTAAAGATACTTTAT 1920
Db 1931 CAAGGACCACTCCACCTGAACCAACACACTCTAAAGAAATATTTTAAAGATACTTTAT 1990
QY 1921 TTTCTTCTTACTCCTTCTCCTCTGATTTTCTTCTTATATTTTCTATCTGTTTTCATC 1980
Db 1991 TTTCTTCTTACTCCTTCTCCTCTGATTTTCTTCTTATATTTTCTATCTGTTTTCATC 2050
QY 1981 TCATTTCCAAAGTCTGACAGACACAGGAAGTCTGCTCATGCGCTTTAGATGAATA 2040
Db 2051 TCATTTCCAAAGTCTGACAGACACAGGAAGTCTGCTCATGCGCTTTAGATGAATA 2110
QY 2041 GAAGTCAAGGTCCTCCTCACTCTAGTCACTAAAGAGATTTTACTCCCCAGCCAGAAA 2100
Db 2111 GAAGTCAAGGTCCTCCTCACTCTAGTCACTAAAGAGATTTTACTCCTCCAGCCAGAAA 2170
QY 2101 GGTGATTTCTCTTTTACCATTTCTGGGACTTTAGTCTTAATTAGTACCTTATTACA 2160
Db 2171 GGTGATTTCT-TCTTTACCATTTCTGGGACTTTAGTCTTAATTAGTACCTTATTACA 2229
QY 2161 GGAATGCTAAGTACCTCTCTGTGGAACAATCTGCAATGTCTAAATGCGCTTAAAGA 2220
Db 2230 GGAATGCTAAGTACCTCTCTGTGGAACAATCTGCAATGTCTAAATGCGCTTAAAGA 2289
QY 2221 GCCCATTTCTTAGCTGCTGAAATCAGTCTCTTTCACCTCTTCAGAGAACAGGGATGT 2280
Db 2290 GCCCATTTCTTAGCTGCTGAAATCAGTCTCTTTCACCTCTTCAGAGAACAGGGATGT 2349
QY 2281 ACCTACCCGCGAGTAGGTAGATGTGGGTGGTGCATGTTAATTTCCCTAGAAAGTTCCA 2340
Db 2350 ACCTACCCGCGAGTAGGTAGATGTGGGTGGTGCATGTTAATTTCCCTAGAAAGTTCCA 2409
QY 2341 AGCCCTGTTCTCGCTAAAGTGTATGTCAGTTCAGAGATGTGTAATGAGCATGG 2400
Db 2410 AGCCCTGTTCTCGCTAAAGTGTATGTCAGTTCAGAGATGTGTAATGAGCATGG 2469
QY 2401 CTGTGTAAGATCAGAGAGCCCACTTGATTTATAGTATAGCCCTTCCTCACTCCACCA 2460
Db 2470 CTGTGTAAGATCAGAGAGCCCACTTGATTTATAGTATAGCCCTTCCTCACTCCACCA 2529
QY 2461 GACTTGCTCATTTTTCGAGTTTAACTAGACTACACTTAT 2503
|||||

Db 2530 GACTTGCTCATTTTTCGAGTTTAACTAGACTACACTTAT 2572
RESULT 4
AAV73924
ID AAV73924 standard; DNA; 3616 BP.
XX AC AAV73924;
XX DT 04-MAR-1999 (first entry)
XX DE Human SAHH DNA #1.
XX KW S-adenosyl-5-homocysteine hydrolase; SAHH; human; drug screening;
KW treatment; infection; cancer; autoimmune disease; detection; diagnosis;
KW gene mapping; antisense; therapy; antagonist; immunoassay; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT CDS 56..1558
FT /*tag= a
FT /product= "SAHH"
XX PN US5854023-A.
XX PD 29-DEC-1998.
XX PF 17-JUL-1997; 97US-0896005.
XX PR 17-JUL-1997; 97US-0896005.
XX PA (INCY-) INCYTE PHARM INC.
XX PI Corley NC, Hillman JL, Lal P, Shah P.
XX DR WPI: 1999-094906/08.
XX DR P-PSDB; AAW90061.
XX PT Nucleic acid encoding human S-adenosyl-5-homocysteine hydrolase -
PT for production of recombinant enzyme, useful for diagnosis,
PT treatment and prevention of cancers, infections and autoimmune
PT diseases
XX PS Disclosure; Fig 1A-I; 40pp; English.
XX CC This sequence encodes a human S-adenosyl-5-homocysteine hydrolase (SAHH).
CC The SAHH protein can be used to generate specific antibodies and in drug
CC screening to identify specific binding agents. Antagonists of the
CC protein are used to treat or prevent a wide range of viral, bacterial,
CC fungal, parasitic, protozoal or helminthic infections, many cancers
CC (leukaemia, lymphoma or solid tumours), and many autoimmune diseases
CC (e.g. acquired immune deficiency syndrome, allergy, asthma, diabetes
CC mellitus, multiple sclerosis etc). All these conditions may be treated by
CC expressing antisense sequences, triplex-forming agents or ribozymes
CC directed against the nucleic acid. The nucleic acid and its fragments can
CC be used as probes or primers for detecting and quantifying gene
CC expression, for diagnosis or monitoring of disease, to identify genetic
CC variations, mutations or polymorphisms, in gene mapping and as antisense
CC therapeutics. Antibodies are used directly as antagonists, indirectly to
CC deliver active agents to SAHH-expressing cells, to diagnose and monitor
CC diseases in standard immunoassays, in competitive drug screens and to
CC isolate the protein from natural sources.
XX SQ Sequence 3616 BP; 1017 A; 782 C; 826 G; 991 T; 0 other;
Query Match 59.4%; Score 1522; DB 20; Length 3616;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1572; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 292 GGGAGAGCTGAAGCAGCGCAAGAGATGAGAGCGCGAGAAGTACTCCTTCATGGCA 351
|||||
Db 3 GGGAGAGCTGAAGCAGCGCAAGAGATGAGAGCGCGAGAAGTACTCCTTCATGGCA 62

```
QY 352 CCGTCACCAAGGCGCCCAAGAACAAATCCAGTTTGTGATGACATGACAGAGTTCACCA 411
|||||
Db 63 CCGTCACCAAGGCGCCCAAGAACAAATCCAGTTTGTGATGACATGACAGAGTTCACCA 122
QY 412 AATCCCCACCAAAACTGGCCGAGATCTTTGTCTCGCTCGATCTCACAGTCTCCACTG 471
|||||
Db 123 AATCCCCACCAAAACTGGCCGAGATCTTTGTCTCGCTCGATCTCACAGTCTCCACTG 182
QY 472 ACAGCTACAGTTCAGCTGCATCTACACAGATAGCTCTGATGATGAGTTTCTCCCGAG 531
|||||
Db 183 ACAGCTACAGTTCAGCTGCATCTACACAGATAGCTCTGATGATGAGTTTCTCCCGAG 242
QY 532 AGAAGCAGCAAAACCAACTCCAAAGGCGACAGCAATTTCTGTGAAGAACATCAAGCAG 591
|||||
Db 243 AGAAGCAGCAAAACCAACTCCAAAGGCGACAGCAATTTCTGTGAAGAACATCAAGCAG 302
QY 592 CAGAATTTGGACGCGCGGAGATGAGATTCAGAGCAACATGTCTGCTGATTTTCAC 651
|||||
Db 303 CAGAATTTGGACGCGCGGAGATGAGATTCAGAGCAACATGTCTGCTGATTTTCAC 362
QY 652 TCAGGAAACGTGCTCAGGGGGGAGAAAGCCCTTGCTGTCTAAATAGTGGCTGTACAC 711
|||||
Db 363 TCAGGAAACGTGCTCAGGGGGGAGAAAGCCCTTGCTGTCTAAATAGTGGCTGTACAC 422
QY 712 ACATCAGACGCCAGACAGCGGTGTGATGAGACACTCTGTGCCCTGGGGGCTCAGTGCC 771
|||||
Db 423 ACATCAGACGCCAGACAGCGGTGTGATGAGACACTCTGTGCCCTGGGGGCTCAGTGCC 482
QY 772 GCTGCTCTGCTGTGTACATCTACTCAACTCAGATGAAGTAGCTGCAGCAGTGGCTGAGG 831
|||||
Db 483 GCTGCTCTGCTGTGTACATCTACTCAACTCAGATGAAGTAGCTGCAGCAGTGGCTGAGG 542
QY 832 CTGAGATTGCACTGTCTGCTGGAAGGGCGAGTCAGAAAGTAGCTCTGCTGCTGATTTG 891
|||||
Db 543 CTGAGATTGCACTGTCTGCTGGAAGGGCGAGTCAGAAAGTAGCTCTGCTGCTGATTTG 602
QY 892 ACCGCTGTGTGAACATGATGGGTGAGGCGCAACATGATCTGTGATGATGGGGAGACT 951
|||||
Db 603 ACCGCTGTGTGAACATGATGGGTGAGGCGCAACATGATCTGTGATGATGGGGAGACT 662
QY 952 TAACCCACTGGGTTTATAAGAGATTAACAACGTGTTTAAAGATCCGAGGCAATGCTG 1011
|||||
Db 663 TAACCCACTGGGTTTATAAGAGATTAACAACGTGTTTAAAGATCCGAGGCAATGCTG 722
QY 1012 AAGAGAGCGTGAAGTGTGTACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTG 1071
|||||
Db 723 AAGAGAGCGTGAAGTGTGTACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTG 782
QY 1072 TTCCGGCCATGAACGTCAATGATCTGTACCAACAGAAAGTTGATTAATCTGTACTGCT 1131
|||||
Db 783 TTCCGGCCATGAACGTCAATGATCTGTACCAACAGAAAGTTGATTAATCTGTACTGCT 842
QY 1132 GCCGAGAAATCCATTTTGGATGGCTGAAGAGGACACAGATGTGATGTTGGTGAAC 1191
|||||
Db 843 GCCGAGAAATCCATTTTGGATGGCTGAAGAGGACACAGATGTGATGTTGGTGAAC 902
QY 1192 AAGTGTGTGTGTGTGCTATGTGAGGTGAGGCAAGGCTGTGCTGTCTCAAGCTC 1251
|||||
Db 903 AAGTGTGTGTGTGTGCTATGTGAGGTGAGGCAAGGCTGTGCTGTCTCAAGCTC 962
QY 1252 TTGAGCAATGTGTACATTAACGAAATGCAACCCCATCTGTGCTGCAGGCGCTGCA 1311
|||||
Db 963 TTGAGCAATGTGTACATTAACGAAATGCAACCCCATCTGTGCTGCAGGCGCTGCA 1022
QY 1312 ATGGGTTCAGGGTGTGTAAGCTAAATGAAGTATCCGGCAAGTGCATGTCTAATACTT 1371
|||||
Db 1023 ATGGGTTCAGGGTGTGTAAGCTAAATGAAGTATCCGGCAAGTGCATGTCTAATACTT 1082
QY 1372 GCACAGGAATAAGATGTAGTGCACAGGAGCACTTGATCGCATGAAAAACAGTTGTA 1431
|||||
Db 1083 GCACAGGAATAAGATGTAGTGCACAGGAGCACTTGATCGCATGAAAAACAGTTGTA 1142
```

```
QY 1432 TCGTATGCAATATGGGCCACTCCAAACACAGAAATCGATGTGACCAGCCTCCGCACTCCGG 1491
|||||
Db 1143 TCGTATGCAATATGGGCCACTCCAAACACAGAAATCGATGTGACCAGCCTCCGCACTCCGG 1202
QY 1492 AGCTGACGTGGGAGCGAGTACGTTCTCAGGTGAGCCATGTCTATCTGGCCAGATGGCAAC 1551
|||||
Db 1203 AGCTGACGTGGGAGCGAGTACGTTCTCAGGTGAGCCATGTCTATCTGGCCAGATGGCAAC 1262
QY 1552 GAGTTGCTCTCCCTGGCAGAGGGTCTGTACTCAATTTGAGCTGCTCCACAGTCCACCT 1611
|||||
Db 1263 GAGTTGCTCTCCCTGGCAGAGGGTCTGTACTCAATTTGAGCTGCTCCACAGTCCACCT 1322
QY 1612 TTGTTCTGTCCATCACAGCCACACACAGGCTTTGGCACTGATAGAACTCTAATAATGCA 1671
|||||
Db 1323 TTGTTCTGTCCATCACAGCCACACACAGGCTTTGGCACTGATAGAACTCTAATAATGCA 1382
QY 1672 CCGAGGGGCGATACAGCAGGATGTGTACTTGTCTCTTAAGAAATGGATGAATACGTTG 1731
|||||
Db 1383 CCGAGGGGCGATACAGCAGGATGTGTACTTGTCTCTTAAGAAATGGATGAATACGTTG 1442
QY 1732 CCAGCTTGCACTGCCATCATTTGATGCCCCACCTTACAGAGCTGACAGATGACCAAGCA 1791
|||||
Db 1443 CCAGCTTGCACTGCCATCATTTGATGCCCCACCTTACAGAGCTGACAGATGACCAAGCA 1502
QY 1792 AATATCTGGGACTCAACAAAATGGGCCATTCAAACCTAATTATTACAGATACTAATGGA 1851
|||||
Db 1503 AATATCTGGGACTCAACAAAATGGGCCATTCAAACCTAATTATTACAGATACTAATGGA 1562
QY 1852 CCATACTACCAAG 1864
|||||
Db 1563 CCATACTACCAAG 1575
```

```
RESULT 5
AAFI6004
ID AAFI6004 standard; cDNA; 721 BP.
XX
AC AAFI6004;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:439.
XX
KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
vulnery; gastrointestinal; nephrotropic; antinfecive; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
wound; infectious disease; ss.
XX
OS Homo sapiens.
XX
PN WO200055174-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US05988.
XX
PR 12-MAR-1999; 99US-0124270.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2000-587513/55.
DR P-PSDB; AAB56801.
XX
PT Prostate cancer associated gene sequences, referred to as prostate
cancer antigens, useful for treatment, prevention, and diagnosis of
disorders such as prostate cancer -
XX
PS Claim 1; Page 917-918; 2338bp; English.
```

XX AAF15566 to AAF16505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytosolic,
CC cardioactive, immunomodulatory, muscular, vulnerary, gastrointestinal,
CC nephrotropic, antiinfective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention.

XX Sequence 721 BP; 157 A; 195 C; 243 G; 119 T; 7 other;

Query Match 22.2%; Score 570; DB 21; Length 721;
Best Local Similarity 99.8%; Pred. No. 5.6e-208;
Matches 620; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 53 CCGTCGCTGTCCGGCTGCTGGCTGCGGACAGACAGAGCCGTGGCCACAGACACTCA 112
DB 56 CCGTCGCTGTCCGGCTGCTGGCTGCGGACAGACAGAGCCGTGGCCACAGACACTCA 115
QY 113 GAAGCCGACGACGCTGACGACGAGGCGCGGACGAGAGGCTGGCGCATCCGCTCGAGG 172
DB 116 GAAGCCGACGACGCTGACGACGAGGCGCGGACGAGAGGCTGGCGCATCCGCTCGAGG 175
QY 173 CGCCGCGCGGACGAGCGCGGCGCGCCAGAGGCGGAAAGAGCGCGGCGCGGCTCAGC 232
DB 176 CGCCGCGCGGACGAGCGCGGCGCGCCAGAGGCGGAAAGAGCGCGGCGCGGCTCAGC 235
QY 233 CGCTGCGCGGCGCGCGCGGGAATGTCATGCTGACGCGCATCCGCTGCGCGGCTCGG 292
DB 236 CGCTGCGCGGCGCGCGCGGGAATGTCATGCTGACGCGCATCCGCTGCGCGGCTCGG 295
QY 293 GGAGGAGCTGAAGCAGCGGCAAGAGATCGAGGACGCCGAGAGTACTCCTTCATGGCCAC 352
DB 296 GGAGGAGCTGAAGCAGCGGCAAGAGATCGAGGACGCCGAGAGTACTCCTTCATGGCCAC 355
QY 353 CGTCACCAAGCGCGCCCAAGACCAATCCAGTTGCTGATGACATGCAAGAGTCAACCA 412
DB 356 CGTCACCAAGCGCGCCCAAGACCAATCCAGTTGCTGATGACATGCAAGAGTCAACCA 415
QY 413 ATTCGCCCAACCAACTGGCGGAAGATCTTGTCTGCTGATCTCAGCTCCTCCACTGA 472
DB 416 ATTCGCCCAACCAACTGGCGGAAGATCTTGTCTGCTGATCTCAGCTCCTCCACTGA 475
QY 473 CAGCTACAGTTGAGCTGATCTTACACAGATAGCTCTGATGATGAGGTTTCTCCCGAGA 532
DB 476 CAGCTACAGTTGAGCTGATCTTACACAGATAGCTCTGATGATGAGGTTTCTCCCGAGA 535
QY 533 GAAGCAGCAACCAACTCCAGAGGCGAGCAGCAATTTCTGTGTAAGAATCAAGCAGGC 592
DB 536 GAAGCAGCAACCAACTCCAGAGGCGAGCAGCAATTTCTGTGTAAGAATCAAGCAGGC 595
QY 593 AGAATTTGAGCGCGCGGAGATGAGATTCAGAGCAAGACATGCTCTGATTTCACT 652
DB 596 AGAATTTGAGCGCGCGGAGATGAGATTCAGAGCAAGACATGCTCTGATTTCACT 655
QY 653 CAGGAACGTCAGGCGGA 673
DB 656 CAGGAACGTCAGGCGGA 676

RESULT 6
AAV73925
ID AAV73925 standard; DNA; 2226 BP.
XX AC AAV73925;
XX

DT 04-MAR-1999 (first entry)
XX Human SAHH DNA #2.
DE
XX S-adenosyl-5-homocysteine hydrolase; SAHH; human; drug screening;
KW treatment; infection; cancer; autoimmune disease; detection; diagnosis;
KW gene mapping; antisense; therapy; antagonist; immunoassay; ss.
XX Homo sapiens.
XX
XX US5854023-A.
XX
XX 29-DEC-1998;
PD
XX 17-JUL-1997; 97US-0896005.
PF
XX 17-JUL-1997; 97US-0896005.
PR
XX 17-JUL-1997; 97US-0896005.
XX
XX (INCY-) INCYTE PHARM INC.
XX Corley NC, Hillman J, Lal P, Shah P;
XX WPI: 1999-094906/08.
XX P-PSDB; AAW90061.
XX
XX Nucleic acid encoding human S-adenosyl-5-homocysteine hydrolase -
PT for production of recombinant enzyme, useful for diagnosis,
PT treatment and prevention of cancers, infections and autoimmune
PT diseases
XX
XX Claim 5; Column 39-42; 40pp; English.

XX This sequence encodes a human S-adenosyl-5-homocysteine hydrolase (SAHH).
CC The SAHH protein can be used to generate specific antibodies and in drug
CC screening to identify specific binding agents. Antagonists of the
CC protein are used to treat or prevent a wide range of viral, bacterial,
CC fungal, parasitic, protozoal or helminthic infections, many cancers
CC (leukaemia, lymphoma or solid tumours), and many autoimmune diseases
CC (e.g. acquired immune deficiency syndrome, allergy, asthma, diabetes
CC mellitus, multiple sclerosis etc). All these conditions may be treated by
CC expressing antisense sequences, triplex-forming agents or ribozymes
CC directed against the nucleic acid. The nucleic acid and its fragments can
CC be used as probes or primers for detecting and quantifying gene
CC expression, for diagnosis or monitoring of disease, to identify genetic
CC variations, mutations or polymorphisms, in gene mapping and as antisense
CC therapeutics. Antibodies are used directly as antagonists, indirectly to
CC deliver active agents to SAHH-expressing cells, to diagnose and monitor
CC diseases in standard immunoassays, in competitive drug screens and to
CC isolate the protein from natural sources.

XX Sequence 2226 BP; 649 A; 455 C; 435 G; 675 T; 12 other;

Query Match 21.6%; Score 553; DB 20; Length 2226;
Best Local Similarity 99.6%; Pred. No. 1.5e-201;
Matches 773; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1728 GTTGCCAGCTTGATCTGCCATCATTTGATGCCCACTTACAGAGCTGACAGATGACCA 1787
DB 1 GTTGCCAGCTTGATCTGCCATCATTTGATGCCCACTTACAGAGCTGACAGATGACCA 60
QY 1788 GCAAAATATCTGGGACTCAACAAAATGGGCCATTCAAACTAATTAATACAGATACTAA 1847
DB 61 GCAAAATATCTGGGACTCAACAAAATGGGCCATTCAAACTAATTAATACAGATACTAA 120
QY 1848 TGGACCATACTACCAAGGACGACGACCTGAACACACACTCTAAAGAAATATTTTAA 1907
DB 121 TGGACCATACTACCAAGGACGACGACCTGAACACACACTCTAAAGAAATATTTTAA 180
QY 1908 ACATAACTTTTATTTCTTCTTACCTTCTCTCTGATTTTTCCTAATATTCATTC 1967
DB 181 ACATAACTTTTATTTCTTCTTACCTTCTCTCTGATTTTTCCTAATATTCATTC 240
QY 1968 TTGTTTTTTCATCTCATTTATCCAGTTCTGCAAGACACACAGGAAGTCTGCTCATGGCTC 2027

|||||
Db 241 TTGTTTTCATTCATATATCCAGTCTGACAGACCACAGAACTGCTCATGGCTC 300
QY 2028 TTATGATGAATAGAGTTCAGGGTCCCTCCTCTCTAGTCACTAAGAAGGATTACTCC 2087
Db 301 TTATGATGAATAGAGTTCAGGGTCCCTCCTCTCTAGTCACTAAGAAGGATTACTCT 360
QY 2088 CCCAGCCCCAGAAAGGTGATCTCTCTTTTACCATTTCTGGGAGCTTTAGTCTTAATTAGG 2147
Db 361 CCCAGCCCCAGAAAGGTGATCTCTCTTTTACCATTTCTGGGAGCTTTAGTCTTAATTAGG 419
QY 2148 TACCTTTATTAACAGAAATGCTAAGTACCTTCTCTGGAACAATCTGCAATGTCTAAA 2207
Db 420 TACCTTTATTAACAGAAATGCTAAGTACCTTCTCTGGAACAATCTGCAATGTCTAAA 479
QY 2208 TCGCCTTAAAGAGCCCATTTCTTAGCTGCTGAATCAGTGTCTTTCACTTCTTCAAGAG 2267
Db 480 TCGCCTTAAAGAGCCCATTTCTTAGCTGCTGAATCAGTGTCTTTCACTTCTTCAAGAG 539
QY 2268 AAGCAGGATGTACCTACCCGGCAGGTAGGTAGATGTGGGTGTCATGTATTTC 2327
Db 540 AAGCAGGATGTACCTACCCGGCAGGTAGGTAGATGTGGGTGTCATGTATTTC 599
QY 2328 CTTAGAGTTCAGACCCCTGTTTCTGCTAAAGGTGTATGTCCAGTCCAGAGATGTGT 2387
Db 600 CTTAGAGTTCAGACCCCTGTTTCTGCTAAAGGTGTATGTCCAGTCCAGAGATGTGT 659
QY 2388 ATATGAGCATGCTGTGTTAAGATCAGAGGCCCACTGTGATTTATAGTATAGCCCTTCC 2447
Db 660 ATATGAGCATGCTGTGTTAAGATCAGAGGCCCACTGTGATTTATAGTATAGCCCTTCC 719
QY 2448 TCCACTCCACAGACTTGTCTCATTTTTCGAGTTTAACTAGACTACACTTATT 2503
Db 720 TCCACTCCACAGACTTGTCTCATTTTTCGAGTTTAACTAGACTACACTTATT 775
RESULT 7
ABV44141/c
ID ABV44141 standard; cDNA; 504 BP.
XX AC ABV44141;
XX DT 16-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 44132.
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN WO200160860-A2.
XX PD 23-AUG-2001.
XX PF 20-FEB-2001; 2001WO-US05171.
XX PR 17-FEB-2000; 2000US-183319P.
XX PR 16-MAR-2000; 2000US-189862P.
XX PR 25-MAY-2000; 2000US-207454P.
XX PR 09-JUN-2000; 2000US-211314P.
XX PR 18-JUL-2000; 2000US-219007P.
XX PR 13-DEC-2000; 2000US-255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege WO, Monahan JE;
XX DR WPI; 2001-662795/76.
XX PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer

XX PS Claim 1; Page 8770; 11750pp; English.
XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX SQ Sequence 504 BP; 117 A; 150 C; 116 G; 121 T; 0 other;
Query Match 16.5%; Score 422; DB 23; Length 504;
Best Local Similarity 100.0%; Pred. No. 1.8e-151;
Matches 422; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 706 GTACACATCACAGCCACAGACGGGTGTGATGAGACACTCTGCCCCGGGGCTC 765
Db 502 GTACACATCACAGCCACAGACGGGTGTGATGAGACACTCTGCCCCGGGGCTC 443
QY 766 AGTCCGCTGTCTGCTTGAACATCTACTCAACTCAGAATGAAGTAGCTGCAGCACTGG 825
Db 442 AGTCCGCTGTCTGCTTGAACATCTACTCAACTCAGAATGAAGTAGCTGCAGCACTGG 383
QY 826 CTGAGGCTGAGTTCAGTGTTCGCTTGAAGGGCGAGTCAGAGATGACTTCTGTGTGT 885
Db 382 CTGAGGCTGAGTTCAGTGTTCGCTTGAAGGGCGAGTCAGAGATGACTTCTGTGTGT 323
QY 886 GTATTGACCGCTGTGTGAACATGATGGGTGGCAGGCCAACATGATCTGTGATGGGG 945
Db 322 GTATTGACCGCTGTGTGAACATGATGGGTGGCAGGCCAACATGATCTGTGATGGGG 263
QY 946 GAGACTTAACCCCACTGGGTTTAAAGAGTATCCAAACGTGTTTAAAGAGATCCGAGGCA 1005
Db 262 GAGACTTAACCCCACTGGGTTTAAAGAGTATCCAAACGTGTTTAAAGAGATCCGAGGCA 203
QY 1006 TTGTGAAGAGAGCGTGAAGTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGC 1065
Db 202 TTGTGAAGAGAGCGTGAAGTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGC 143
QY 1066 TCTGTGTTCCGGCCATGAACGTCAATGATCTGTATCCAAACAGAGATTGATACTGT 1125
Db 142 TCTGTGTTCCGGCCATGAACGTCAATGATCTGTATCCAAACAGAGATTGATACTGT 83
QY 1126 AC 1127
Db 82 AC 81
RESULT 8
AAZ80766/c
ID AAZ80766 standard; cDNA; 636 BP.
XX AC AAZ80766;
XX DT 07-APR-2000 (first entry)
XX DE Human colon cancer cell line SW480 cDNA clone SEQ ID NO:850.
XX KW Human; gene expression product; diagnosis; tumour; colon cancer;
KW colorectal adenocarcinoma; cell line SW480; cell proliferation;
KW cytostatic; sarcoma; breast cancer; neoplasia; dysplasia;
KW hyperplasia; ds.


```
XX Claim 1; Fig 1; 299pp; English.
PS
XX
CC The present invention describes an isolated polypeptide comprising an
CC immunogenic portion of an ovarian carcinoma protein (or its variants).
CC Ovarian carcinoma proteins, and polynucleotides encoding them, have
CC cytostatic activity and can be used in gene therapy and vaccines.
CC Ovarian carcinoma polypeptides, nucleic acids, antibodies and vaccines
CC are useful for the prevention, diagnosis and treatment of cancer,
CC preferably ovarian cancer. AAA69691 to AAA70077 and AAB12552 to AAB12557
CC represent human ovarian carcinoma polynucleotides and proteins used in
CC the exemplification of the present invention.
XX
SQ Sequence 292 BP; 98 A; 68 C; 60 G; 66 T; 0 other;

Query Match          9.2%; Score 237; DB 21; Length 292;
Best Local Similarity 100.0%; Pred. No. 7.8e-81;
Matches 237; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2267 GAAGCAGGATGTTACCTACCCGGCAGGTAGTTAGATGTGGTGTGCATGTTAATTTC 2326
    |||||||
DB 276 GAAGCAGGATGTTACCTACCCGGCAGGTAGTTAGATGTGGTGTGCATGTTAATTTC 217
    |||||||

QY 2327 CCTTAGAGTTCCAAAGCCCTGTTCCCTGCGTAAAGTGTGTATGTCAGTTCAGAGATGTG 2386
    |||||||
DB 216 CCTTAGAGTTCCAAAGCCCTGTTCCCTGCGTAAAGTGTGTATGTCAGTTCAGAGATGTG 157
    |||||||

QY 2387 TATAATGAGCATGGCTTGTTAAGATCAGAGGCCCACTTGATTATAGTATAGCCCTTC 2446
    |||||||
DB 156 TATAATGAGCATGGCTTGTTAAGATCAGAGGCCCACTTGATTATAGTATAGCCCTTC 97
    |||||||

QY 2447 CTCACACTCCACACAGACTTGCTCATTTTTCGAGTTTAACTAGACTACACTCTATT 2503
    |||||||
DB 96 CTCACACTCCACACAGACTTGCTCATTTTTCGAGTTTAACTAGACTACACTCTATT 40
    |||||||

RESULT 12
ABN72624/c
ID ABN72624 standard; DNA; 292 BP.
XX
XX ABN72624;
AC
XX 02-JUL-2002 (first entry)
DT
XX Ovarian carcinoma antigen polynucleotide #40.
DE
XX Human; immunostimulant; cytostatic; cancer; ovarian carcinoma; ds.
KW
XX Homo sapiens.
OS
XX WO200206317-A2.
PN
XX 24-JAN-2002.
PD
XX
XX 17-JUL-2001; 2001WO-US22635.
PF
XX
XX 17-JUL-2000; 2000US-0617747.
PR 10-AUG-2000; 2000US-0636801.
PR 20-SEP-2000; 2000US-0667857.
PR 04-APR-2001; 2001US-0827271.
PR 18-JUN-2001; 2001US-0884441.
XX
XX (CORI-) CORIXA CORP.
PA
XX
XX Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
PI Reed SG, Vedvick TS, Carter D, Hill P, Albone E;
XX
XX WPI; 2002-164781/21.
XX
XX Polypeptides comprising an immunogenic portion of an ovarian carcinoma
XX protein or its variants, useful for stimulating an immune response in a
XX patient and treating ovarian cancer -
```

```
PS Example 1; Page 104; 408pp; English.
XX
XX This invention relates to polypeptides comprising an immunogenic
XX portion of an ovarian carcinoma protein which acts as an
XX immunostimulant and is cytostatic. The polypeptides, polynucleotides,
XX antibodies, fusion proteins, T cell populations and antigen presenting
XX cells that express the polypeptides are useful for stimulating an
XX immune response in a patient and treating ovarian cancer. This
XX sequence represents DNA related to the invention.
XX
SQ Sequence 292 BP; 98 A; 68 C; 60 G; 66 T; 0 other;

Query Match          9.2%; Score 237; DB 24; Length 292;
Best Local Similarity 100.0%; Pred. No. 7.8e-81;
Matches 237; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2267 GAAGCAGGATGTTACCTACCCGGCAGGTAGTTAGATGTGGTGTGCATGTTAATTTC 2326
    |||||||
DB 276 GAAGCAGGATGTTACCTACCCGGCAGGTAGTTAGATGTGGTGTGCATGTTAATTTC 217
    |||||||

QY 2327 CCTTAGAGTTCCAAAGCCCTGTTCCCTGCGTAAAGTGTGTATGTCAGTTCAGAGATGTG 2386
    |||||||
DB 216 CCTTAGAGTTCCAAAGCCCTGTTCCCTGCGTAAAGTGTGTATGTCAGTTCAGAGATGTG 157
    |||||||

QY 2387 TATAATGAGCATGGCTTGTTAAGATCAGAGGCCCACTTGATTATAGTATAGCCCTTC 2446
    |||||||
DB 156 TATAATGAGCATGGCTTGTTAAGATCAGAGGCCCACTTGATTATAGTATAGCCCTTC 97
    |||||||

QY 2447 CTCACACTCCACACAGACTTGCTCATTTTTCGAGTTTAACTAGACTACACTCTATT 2503
    |||||||
DB 96 CTCACACTCCACACAGACTTGCTCATTTTTCGAGTTTAACTAGACTACACTCTATT 40
    |||||||

RESULT 13
ABL79045
ID ABL79045 standard; cDNA; 553 BP.
XX
XX ABL79045;
AC
XX 17-MAY-2002 (first entry)
DT
XX Human ovarian cancer related cDNA clone SEQ ID NO:2023.
DE
XX Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
KW
XX Homo sapiens.
OS
XX WO200192581-A2.
PN
XX 06-DEC-2001.
PD
XX
XX 29-MAY-2001; 2001WO-US17756.
PF
XX
XX 26-MAY-2000; 2000US-207484P.
PR
XX
XX (CORI-) CORIXA CORP.
PA
XX
XX Algate PA, Harlocker SL, Jones R;
PI
XX WPI; 2002-122075/16.
XX
XX Composition for therapy and diagnosis of ovarian cancer comprising
XX polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
XX polypeptide, antibody specific to polypeptide or T cell expressing
XX polypeptide
XX
XX Claim 1; SEQ ID 2023; 489pp; English.
XX
XX The present invention describes a composition (I) comprising: carriers
XX and immunostimulants; and a polypeptide (II) of a ovarian tumour
XX polypeptide encoded by a polynucleotide (III) having a cDNA sequence
XX (SI) from the 10912 nucleotide sequences as given in ABL77023 to
XX ABL87934, (III) encoding (II) having a sequence (S2), a T cell
```

CC population of (II), or antigen presenting cells that express (II).
CC (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to
CC (S1) can be used for detecting ovarian cancer in a patient's biological
CC sample preferably serum or ovarian tissue. The method comprises
CC contacting a biological sample from a patient with (IV), detecting the
CC amount of polynucleotide hybridising to (IV) and comparing the amount to
CC a predetermined cutoff value and thereby detecting ovarian cancer in the
CC patient, where the amount of polynucleotide hybridising to (IV) is
CC detected preferably by polymerase chain reaction (PCR). (I) comprising
CC (III) and/or (II) is useful for stimulating and/or expanding T cells
CC specific for an ovarian tumour protein comprising contacting T cells
CC with (III) or (II). (III) is useful in design and preparation of
CC ribozyme molecules for inhibiting expression of the tumour polypeptides
CC and proteins in tumour cells; and to isolate a full length gene from a
CC suitable library e.g., a tumour cDNA library using well known
CC techniques.

SO Sequence 553 BP; 159 A; 136 C; 128 G; 127 T; 3 other;

Query Match 8.7%; Score 224; DB 24; Length 553;
Best Local Similarity 100.0%; Pred. No. 6.5e-76;
Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1322 GGTGGTAAAGCTAATGAGTCATCCGGCAAGTCGATGTCGTAATACTGCACAGGAAA 1381
DB 1 GGTGGTAAAGCTAATGAGTCATCCGGCAAGTCGATGTCGTAATACTGCACAGGAAA 60
OY 1382 TAAGATGTAGTGACACGGGAGCAGCTGGATGCGATGAAAAACAGTTGATCGTATGCAA 1441
DB 61 TAAGATGTAGTGACACGGGAGCAGCTGGATGCGATGAAAAACAGTTGATCGTATGCAA 120
OY 1442 TATGGGCCACTCCACACAGAAATGATGTGACCGCTCCGACCTCCGGAGCTGACGTG 1501
DB 121 TATGGGCCACTCCACACAGAAATGATGTGACCGCTCCGACCTCCGGAGCTGACGTG 180
OY 1502 GGAGCGAGTACGTCTCAGGTGACCAATGTCATCTGGCCAGATG 1545
DB 181 GGAGCGAGTACGTCTCAGGTGACCAATGTCATCTGGCCAGATG 224

RESULT 14:
AAV06354/c
ID AAV06354 standard; DNA; 1877 BP.

XX AAV06354;
AC
XX
DT 30-APR-1998 (first entry)
XX
DE AL-2-long (AL-21) protein encoding DNA.

XX
KW AL-21; AL-2; AL-2-long; human; treatment; neurological disorder; tumour;
KW rheumatoid arthritis; wound healing; paralysis; angiogenesis; leukaemia;
KW psoriasis; Alzheimer's disease; epilepsy; ds.

XX Homo sapiens.

XX
FH Key Location/Qualifiers
FT CDS 244..1611
FT /*tag= a
FT /product= "AL-21"
FT 244..321
FT /*tag= b
FT mat_peptide 322..1608
FT /*tag= c

XX MO9740153-A1.
XX
XX 30-OCT-1997.
XX
XX 17-APR-1997; 97WO-US06345.
XX
XX 19-APR-1996; 96US-0635130.
XX

PA (GETH) GENENTECH INC.
XX
XX Caras IW;
XX
DR WPI; 1997-535837/49.
DR P-PSDB; AAW33698.
XX
PT Human AL-2 neurotrophic factor and related DNA - used to develop
PT products for, e.g. treating neurologic disorders, angiogenesis
PT disorders, tumours or rheumatoid arthritis or for wound healing
XX
PS Claim 3; Fig 1A-C; 86pp; English.

XX
CC This DNA encodes a AL-2-long (AL-21) protein. AL-2 is a novel Eph-related
CC tyrosine kinase receptor ligand. AL-2 can be administered to patients in
CC whom the nervous system has been damaged by trauma, surgery, stroke,
CC ischaemia, infection, metabolic disease, nutritional deficiency,
CC malignancy, or toxic agents, to promote the survival or growth of
CC neurons. They can be used to treat motoneuron disorders such as
CC amyotrophic lateral sclerosis (Lou Gehrig's disease), Bell's palsy, and
CC various conditions involving spinal muscular atrophy, or paralysis. AL-2
CC can be used to treat human neurodegenerative disorders, such as
CC Alzheimer's disease, Parkinson's disease, epilepsy, demyelinating
CC diseases such as multiple sclerosis, Huntingtons chorea, Down's syndrome,
CC nerve deafness, Menier's disease, and other disorders of the cerebellum.
CC AL-2 can be used as cognitive enhancer, to enhance learning particularly
CC in dementias or trauma, since they can promote axonal outgrowth and
CC synaptic plasticity, particularly of hippocampal neurons that express
CC AL-2. AL-2 can also be used for wound healing, i.e. accelerating
CC neovascularisation of, e.g. burns and ulcers. The encoding nucleic acids
CC are useful in preparing antibodies that specifically bind to the AL-2
CC protein. The antibodies and the AL-2 antagonists are useful in diagnosing
CC and treating various neuronal disorders. AL-2 antagonists can be used
CC for modulating angiogenesis. They can also be used for the treatment of
CC tumours, acute myeloid leukaemia (AML), chronic myeloid leukaemia (CML),
CC myelodysplastic syndrome (MDS), diabetic retinopathy, neovascular
CC glaucoma, psoriasis and rheumatoid arthritis.

XX
SQ Sequence 1877 BP; 334 A; 525 C; 615 G; 401 T; 2 other;

Query Match 7.4%; Score 190; DB 18; Length 1877;
Best Local Similarity 100.0%; Pred. No. 5.1e-63;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1283 CCCCATCTGTGCTGTGACAGGCCCTGCATGATGGGTTGAGGTTAAAGCTAAATGAGT 1342
DB 1446 CCCCATCTGTGCTGTGACAGGCCCTGCATGATGGGTTGAGGTTAAAGCTAAATGAGT 1387
OY 1343 CATCCGCAAGTGTGCTGTAATACTGCACAGGAAATGAATGTAGTACACGGGA 1402
DB 1386 CATCCGCAAGTGTGCTGTAATACTGCACAGGAAATGAATGTAGTACACGGGA 1327
OY 1403 GCACTTGATGCGATGAAACAGAGTTGATGCTATGCAATATGGGCCACTCCAACACAGA 1462
DB 1326 GCACTTGATGCGATGAAACAGAGTTGATGCTATGCAATATGGGCCACTCCAACACAGA 1267
OY 1463 AATCGATGTG 1472
DB 1266 AATCGATGTG 1257

RESULT 15
*ABN93848
ID ABN93848 standard; DNA; 384 BP.

XX ABN93848;
XX
XX 13-AUG-2002 (first entry)
XX
XX Gene #346 used to diagnose liver cancer.
XX
XX Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
XX

KW metastatic liver tumour; cytostatic; expression profile; disease state;
KM disease progression; drug toxicity; drug efficacy; drug metabolism.
XX Homo sapiens.
OS WO200229103-A2.
PN 11-APR-2002.
XX
PD 02-OCT-2001; 2001WO-US30589.
XX
PF 02-OCT-2000; 2000US-237054P.
XX
PR (GENE-) GENE LOGIC INC.
XX
PA Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX WPI; 2002-426119/45.
DR
XX
XX
PT Diagnosing and detecting the progression of liver cancer,
PT hepatocellular carcinoma or metastatic liver tumor in a patient,
PT involves detecting the level of expression of two or more genes in a
PT liver tissue sample
XX
XX
PS Claim 1; SEQ ID NO 346; 298bp; English.
XX
CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytosstatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 384 BP; 95 A; 77 C; 113 G; 99 T; 0 other;

Query Match 6.6%; Score 168; DB 24; Length 384;
Best Local Similarity 100.0%; Pred. No. 1.6e-54;
Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1065 CTCGTCTCCGCCATGACGTCAATGATTCTGTACCAACAGAGTTTGATACTTG 1124
Db 73 CTCGTCTCCGCCATGACGTCAATGATTCTGTACCAACAGAGTTTGATACTTG 132

QY 1125 TACTGCTGCCGAGATTCATTGATGCGCTGAAGAGACACAGATGTGATGTTTGGT 1184
Db 133 TACTGCTGCCGAGATTCATTGATGCGCTGAAGAGACACAGATGTGATGTTTGGT 192

QY 1185 GGGAAACAAGTGTGTGTGGCTATGAGGTAGGCAAGGGCTGC 1232
Db 193 GGGAAACAAGTGTGTGTGGCTATGAGGTAGGCAAGGGCTGC 240

Search completed: April 20, 2003, 19:20:50
Job time : 561 secs

JOURNAL Patent: WO 9814562-A 1 09-APR-1998;
HART DEREK JOHN (NZ)
FEATURES Location/Qualifiers
source 1. 2563
/organism="Homo sapiens"
/db_xref="taxon:9606"
CDS
<3. 1847
/note="Open reading frame extends without a stop codon for the full 5' nucleotide sequence. The initiation codon has yet to be identified."
/codon_start=1
/protein_id="CAC09285.1"
/db_xref="GI:10190061"
/translation="RGVGARSCCFWFGHRCPPALGCRDPAKATAPOKPTOLDA
GAGRRVDRVSEGAARAGRAPEGERGGGSAAGAGGMSMPDAMPVGEELKQ
AKEIEDAEKSEMAVTTKAPKKIOFADDMQEFKPEPTKGRSLRSISOSTDYS
SAASYTDSDEVSFPEKQOTNSKSSNECVKNKOAEGREIEIAEODMSALSLR
KRAOGEKPLAGAKIVGCTHIAOTAVLIETLCAIACOCRSACNISTONEVAALAE
AGVAEFAWKSEDEDFWMCIDRCVNDWQANMILDDGDLTHVYKKYPNVFKIRG
IVEESYTGVRHLYOLSKAGKLCVPAMNVNDSYTKQFDNLCCRESILDGKRTDVM
FGKQVVVCGYGEVKGCCCAALKALGAIVYITEIDPICALQACMDGERVVKLNEVIRQ
VDVITCTGKNKVVTRHLDRLNSCSTVPTFVLSITATQALALIELYNAPEGRYKQDVY
HVIMPDKRVVLAEGRLNLSCSTVPTFVLSITATQALALIELYNAPEGRYKQDVY
LIPKMEVYASLHLPFSDAHLTELDQAKYLGKNGPFPNYYRY"

BASE COUNT 646 a 604 c 677 g 636 t
ORIGIN

Query Match 100.0%; Score 2563; DB 6; Length 2563;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGCGGGGAGTGGAGCTGGAGCTGCTGCTTCTGTTCTCTTGCGCCACCGTGGCT 60
DB 1 GGGCGGGGAGTGGAGCTGGAGCTGCTGCTTCTGTTCTCTTGCGCCACCGTGGCT 60
OY 61 GTCCGGCTGCTTGGGCTGCCAGACAGAGCGGTGGCCACAGACTCAGAAAGCCGA 120
DB 61 GTCCGGCTGCTTGGGCTGCCAGACAGAGCGGTGGCCACAGACTCAGAAAGCCGA 120
OY 121 CGCAGCTCGACGAGGCGCGGAGAGGCTGGCGATCGCTGTCGAGAGGCGCGCGC 180
DB 121 CGCAGCTCGACGAGGCGCGGAGAGGCTGGCGATCGCTGTCGAGAGGCGCGCGC 180
OY 181 GGGCAGGCGGGCGGCGCCAGAGGGGAAAGAGCGGGGCGGGGTCAAGCCGCTGGCC 240
DB 181 GGGCAGGCGGGCGGCGCCAGAGGGGAAAGAGCGGGGCGGGGTCAAGCCGCTGGCC 240
OY 241 GGGCGGGCGGGGGAATGTCATGCTGACCGCGATGCCGCTGCCGGGTGGGAGGAGC 300
DB 241 GGGCGGGCGGGGGAATGTCATGCTGACCGCGATGCCGCTGCCGGGTGGGAGGAGC 300
OY 301 TGAAGCAGGCGCAAGAGATCGAGAGCGCCGAGAAGTACTCCTTCATGGCCACCGTCACCA 360
DB 301 TGAAGCAGGCGCAAGAGATCGAGAGCGCCGAGAAGTACTCCTTCATGGCCACCGTCACCA 360
OY 361 AGCGCGCCAGAGCAATTCAGTTGCTGATGACATGACAGAGTTCACCAATTCGCCA 420
DB 361 AGCGCGCCAGAGCAATTCAGTTGCTGATGACATGACAGAGTTCACCAATTCGCCA 420
OY 421 CCAAACTGGCCGAGAGATCTTGTCTGCTGATCTCACAGTCTCCACTGACAGCTACA 480
DB 421 CCAAACTGGCCGAGAGATCTTGTCTGCTGATCTCACAGTCTCCACTGACAGCTACA 480
OY 481 GTTCAGCTGCATCTACACAGATAGCTCTGATGATGAGGTTTCTCCCGAGAGAGCAGC 540
DB 481 GTTCAGCTGCATCTACACAGATAGCTCTGATGATGAGGTTTCTCCCGAGAGAGCAGC 540
OY 541 AAACCACTCCAAGGCGAGCAATTTCTGTGTAAGAACATCAAGCAGAGAAATTG 600
DB 541 AAACCACTCCAAGGCGAGCAATTTCTGTGTAAGAACATCAAGCAGAGAAATTG 600
OY 601 GACGCCGGAGATTGAGATTGACAGCAGACATGTCTGCTCTGATTTCACTCAGGAAC 660
DB 601 GACGCCGGAGATTGAGATTGACAGCAGACATGTCTGCTCTGATTTCACTCAGGAAC 660

DB 601 GACGCCGGAGATTGAGATTGACAGCAGACATGTCTGCTCTGATTTCACTCAGGAAC 660
OY 661 GTGCTCAGGGGAGAGAGCCCTTGGCTGCTGCTAAATAGTGGGCTGTACACATCAGAG 720
DB 661 GTGCTCAGGGGAGAGAGCCCTTGGCTGCTGCTAAATAGTGGGCTGTACACATCAGAG 720
OY 721 CCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGGGGGCTCAGTGGCGCTGCTG 780
DB 721 CCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGGGGGCTCAGTGGCGCTGCTG 780
OY 781 CTTGTACATCTACTCAACTCAGAAATGAGTAGCTGCAGACACTGGCTGAGGCTGAGTTG 840
DB 781 CTTGTACATCTACTCAACTCAGAAATGAGTAGCTGCAGACACTGGCTGAGGCTGAGTTG 840
OY 841 CAGTGTTCGCTTGGAGGGCGAGTCAGAGATGACTTCTGTGTGTATGACCCGCTGTG 900
DB 841 CAGTGTTCGCTTGGAGGGCGAGTCAGAGATGACTTCTGTGTGTATGACCCGCTGTG 900
OY 901 TGAACATGATGGGTGGCAGGCCCAACATGATCCTGATGATGGGGAGACTTAAACCACT 960
DB 901 TGAACATGATGGGTGGCAGGCCCAACATGATCCTGATGATGGGGAGACTTAAACCACT 960
OY 961 GGGTTTATAGAGATATCCAAACGTTTAAAGAGATCCGAGGCAATGTGAGAGAGCG 1020
DB 961 GGGTTTATAGAGATATCCAAACGTTTAAAGAGATCCGAGGCAATGTGAGAGAGCG 1020
OY 1021 TGACTGCTGTTACAGGCTGATACAGCTCTCCAAAGCTGGGAGGCTGTGTTCCGGCCA 1080
DB 1021 TGACTGCTGTTACAGGCTGATACAGCTCTCCAAAGCTGGGAGGCTGTGTTCCGGCCA 1080
OY 1081 TGAAGTCAATGATTTCTGTACCAAGAGAGTTGATTAAGTCTGATGCTGCTCCGAGAAT 1140
DB 1081 TGAAGTCAATGATTTCTGTACCAAGAGAGTTGATTAAGTCTGATGCTGCTCCGAGAAT 1140
OY 1141 CCAATTTGGATGGCTGAAGAGACACAGATGTGATGTTGGTGGAAAGAGTGGTG 1200
DB 1141 CCAATTTGGATGGCTGAAGAGACACAGATGTGATGTTGGTGGAAAGAGTGGTG 1200
OY 1201 TGTGTGCTATGCTGAGGTAGGCAAGGCTGCTGTGCTCTCAAGGCTTTGGAGCAA 1260
DB 1201 TGTGTGCTATGCTGAGGTAGGCAAGGCTGCTGTGCTCTCAAGGCTTTGGAGCAA 1260
OY 1261 TTGTCTACATTAACGAATCGAAGGCTGCTGCTGCTGAGGCTGATGATGGTTCA 1320
DB 1261 TTGTCTACATTAACGAATCGAAGGCTGCTGCTGCTGAGGCTGATGATGGTTCA 1320
OY 1321 GGGTGTAAAGCTAAATGAAGTCAATCCGCGAAGTGTGATGCTTAATACCTGCAGGAA 1380
DB 1321 GGGTGTAAAGCTAAATGAAGTCAATCCGCGAAGTGTGATGCTTAATACCTGCAGGAA 1380
OY 1381 ATAGAATGTAGTGACACGAGGAGCACTGATGATGATGAAAAACAGTTGATGATGCA 1440
DB 1381 ATAGAATGTAGTGACACGAGGAGCACTGATGATGATGAAAAACAGTTGATGATGCA 1440
OY 1441 ATATGGCCCACTCCACACAGAGAAATGATGTGACAGCCCTCCGACTCCGAGCTGACGT 1500
DB 1441 ATATGGCCCACTCCACACAGAGAAATGATGTGACAGCCCTCCGACTCCGAGCTGACGT 1500
OY 1501 GGGAGCGAGTACGTTCTCAGGTGACCATGTCTGCTGCGCAGATGGCAACAGATTGCTC 1560
DB 1501 GGGAGCGAGTACGTTCTCAGGTGACCATGTCTGCTGCGCAGATGGCAACAGATTGCTC 1560
OY 1561 TCCGTGAGAGGGGTGCTACTCAATTTGAGCTGCTCCACAGTTCACACTTGTCTGT 1620
DB 1561 TCCGTGAGAGGGGTGCTACTCAATTTGAGCTGCTCCACAGTTCACACTTGTCTGT 1620
OY 1621 CCATCAGACCAACACAGGCTTGGCACTGATGAACTCTAATATGACCCGAGGGGC 1680
DB 1621 CCATCAGACCAACACAGGCTTGGCACTGATGAACTCTAATATGACCCGAGGGGC 1680
OY 1681 GATACAGCAGAGATGTACTTGTCTCTTAAGAAATGATGATGATGATGATGATGATGAT 1740
DB 1681 GATACAGCAGAGATGTACTTGTCTCTTAAGAAATGATGATGATGATGATGATGATGAT 1740

QY 1741 ATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCACGAAATATCTGG 1800
Db 1741 ATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCACGAAATATCTGG 1800
QY 1801 GACTCAACAAAATGGGCCATTCAAACCTTAATTATACAGATCTAATGACCATACTAC 1860
Db 1801 GACTCAACAAAATGGGCCATTCAAACCTTAATTATACAGATCTAATGACCATACTAC 1860
QY 1861 CAAGACACAGTCCACCTGAAACACACACACTTAAGAAATATTTTAAAGATACTTAT 1920
Db 1861 CAAGACACAGTCCACCTGAAACACACACACTTAAGAAATATTTTAAAGATACTTAT 1920
QY 1921 TTTCTTCTTACTCTCTTCTCTCTGATTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1980
Db 1921 TTTCTTCTTACTCTCTTCTCTCTGATTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1980
QY 1981 TCATTAACCAAGTCTGACAGACACACAGAACTGCTTCAATGCTCTTGAATGAATA 2040
Db 1981 TCATTAACCAAGTCTGACAGACACACAGAACTGCTTCAATGCTCTTGAATGAATA 2040
QY 2041 GAAGTTCAGGGTCCCTCACTCTAGTCACTAAGAAAGATTTTACTCCCCAGCCAGAAA 2100
Db 2041 GAAGTTCAGGGTCCCTCACTCTAGTCACTAAGAAAGATTTTACTCCCCAGCCAGAAA 2100
QY 2101 GGTGATCTCTCTCTTACCATTCTGGGGACTTTAGTCTTAATTAGGTACCTTATACA 2160
Db 2101 GGTGATCTCTCTCTTACCATTCTGGGGACTTTAGTCTTAATTAGGTACCTTATACA 2160
QY 2161 GGAATGCTAAGGTACCTCTCTGTGGAACAATCTGCAATGCTTAATGCTTAAAGA 2220
Db 2161 GGAATGCTAAGGTACCTCTCTGTGGAACAATCTGCAATGCTTAATGCTTAAAGA 2220
QY 2221 GCCCATTTCTTACGCTGCTGAATCAGTCTCTTCTTCTTCTTCTTCTTCTTCTTCT 2280
Db 2221 GCCCATTTCTTACGCTGCTGAATCAGTCTCTTCTTCTTCTTCTTCTTCTTCTTCT 2280
QY 2281 ACCTACCCGGCAGGTAGGTAGATGTGGGTGTCATGTTAATTCCCTTAGAAGTCCA 2340
Db 2281 ACCTACCCGGCAGGTAGGTAGATGTGGGTGTCATGTTAATTCCCTTAGAAGTCCA 2340
QY 2341 AGCCCTGTTTCTGCTGTAAGGTGATGTGTCAGTCTCAAGATGTATATATGAGCATGG 2400
Db 2341 AGCCCTGTTTCTGCTGTAAGGTGATGTGTCAGTCTCAAGATGTATATATGAGCATGG 2400
QY 2401 CTGTTAAGATCAGAGGCCCACTTGATTTATAGTATAGCCCTTCTCCACTCCACCA 2460
Db 2401 CTGTTAAGATCAGAGGCCCACTTGATTTATAGTATAGCCCTTCTCCACTCCACCA 2460
QY 2461 GACTTGCCTATTTTTCAGAGTTTAACTAGACTACACTCTAATTGAGTTAATTGTCC 2520
Db 2461 GACTTGCCTATTTTTCAGAGTTTAACTAGACTACACTCTAATTGAGTTAATTGTCC 2520
QY 2521 TCTAGAGTTTATTTCTGTGTCCAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTT 2563
Db 2521 TCTAGAGTTTATTTCTGTGTCCAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTT 2563

RESULT 2
AF315687 2677 bp mRNA linear PRI 22-MAR-2002
LOCUS AF315687 Homo sapiens S-adenosylhomocysteine hydrolase-like protein mRNA, complete cds.
DEFINITION AF315687
ACCESSION AF315687.1 GI:16588686
VERSION AF315687
KEYWORDS Homo sapiens.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2677)
AUTHORS Dekker, J.W., Budhia, S., Angel, N.Z., Cooper, B.J., Clark, G.J., Hart, D.N. and Kato, M.

TITLE Identification of an S-adenosylhomocysteine hydrolase-like transcript induced during dendritic cell differentiation
JOURNAL Immunogenetics 53 (12), 993-1001 (2002)
MEDLINE 21901265
PUBMED 11904675
REFERENCE 2 (bases 1 to 2677)
AUTHORS Dekker, J.W., Budhia, S., Angel, N.Z., Cooper, B.J., Clark, G.J., Hart, D.N.J. and Kato, M.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-2000) Dendritic Cell Research, Mater Medical Research Institute, Level 3, Aubigny Place, South Brisbane, Queensland 4101, Australia
LOCATION/Qualifiers
FEATURES
source 1. 2677
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p21"
/clone="211(1)B"
/cell_line="Hodgkin's disease cell line L428"
misc_feature 1. 114
/note="derived from 5' RACE"
369. 1961
/note="UCAL; dendritic cell expressed AHcy-like protein; expressed in dendritic cells in blood, langerhans cells and dermal dendritic cells; mRNA increased during monocyte differentiation to monocyte-derived dendritic cells in the presence of GM-CSF and IL-4"
/codon_start=1
/product="S-adenosylhomocysteine hydrolase-like protein"
/protein_id="AAL26869.1"
/db_xref="GI:16588687"
translation="MSMPDAMLPVGEELKAKEIEDAEKYSFMATVTKAPKQIOFADMOEFTKPTKTRSLRSISQSTDSYSSAASYTDSDEVPREKQNTSGS SNECVKNIKAEGRREIEIAEODMSALSLRKAQGEKPLAGAKIVGCTHITATVY LIETICALGAOCRWASACNISTONEVAALAEAVFAWKGESEDFWCCIDRCVAM DGOANMILDDGDLTHWYKKYPNVEFKIRGIVESVTGVHRLXQLSKAGKICVPAM NVNDSVTKOREPNDLYCCRESILDLKRTTDMFGKQVVCYGEVKGCCALKALG AIVYTEIDPICALOACMDGRVVKLNEVIRVDVITCTGNKNVTRHLDRLMNSC IVCNMGHSNTEIDVTSIRTELTWERVRSVDHYTPDGRVLLAEGLNLNLSCTV PFEVLSTATTQALALIELYNAPGRYKQDVYLLPKMDEVYASLHLPFDAHLTELT DDQAKYILGNKNBPENRYR"
BASE COUNT 657 a 648 c 708 g 664 t
ORIGIN
Query Match 100.0%; Score 2563; DB 9; Length 2677;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCCGCGGCGAGTCCGAGCTCGGAGCTGCTCTCTGTTCTCTGTGGCCACCGTCGCT 60
Db 115 GCGCGGCGGCGAGTCCGAGCTCGGAGCTGCTCTCTGTTCTCTGTGGCCACCGTCGCT 174
QY 61 GTCCGGCTGCTTGGGCTGCGAACAACAAGCGGTGGCCACAGCACTCAGAAAGCCGA 120
Db 175 GTCCGGCTGCTTGGGCTGCGAACAACAAGCGGTGGCCACAGCACTCAGAAAGCCGA 234
QY 121 CGCAGCTCGACGCGAGGGGCGGCGAGAGGGTGGCGATGCGGTGTCGGAAGGGCGCGC 180
Db 235 CGCAGCTCGACGCGAGGGGCGGCGAGAGGGTGGCGATGCGGTGTCGGAAGGGCGCGC 294
QY 181 GGGCAGCGCGGCGGCGGCGCAGAGGGGGAAGAAGCGGGGCGGCGGCGGCTGCGCC 240
Db 295 GGGCAGCGCGGCGGCGGCGCAGAGGGGGAAGAAGCGGGGCGGCGGCGGCTGCGCC 354
QY 241 GGGCGCGGCGGCGGGAATGTCATGCTGACCGCATGCCGCTGCCCGGGTCCGGGAGAGC 300
Db 355 GGGCGCGGCGGCGGGAATGTCATGCTGACCGCATGCCGCTGCCCGGGTCCGGGAGAGC 414
QY 301 TGAACAGGCGCAAGAGATCGAGAGCGCGGAGAGTACTCTTCTCATGGCCACCGTCACCA 360
Db 415 TGAACAGGCGCAAGAGATCGAGAGCGCGGAGAGTACTCTTCTCATGGCCACCGTCACCA 474

QY	361	AGCGCCCAAGAACCAATCCAGTTTGTCTGATGACATGCAGAGTTCAACCAATTCCCA	420
Db	475	AGGCGCCCAAGAACCAATCCAGTTTGTCTGATGACATGCAGAGTTCAACCAATTCCCA	534
QY	421	CCAAACTGGCCGAAGATCTTGTCTGCTGATCTCACAGTCTCCATGACAGCTACA	480
Db	535	CCAAACTGGCCGAAGATCTTGTCTGCTGATCTCACAGTCTCCATGACAGCTACA	594
QY	481	GTTCACTGCATCTTACACAGATAGCTCTGATGATGAGTTTCTCCCGAGAGAAGCAGC	540
Db	595	GTTCACTGCATCTTACACAGATAGCTCTGATGATGAGTTTCTCCCGAGAGAAGCAGC	654
QY	541	AAACCACTCCAAGGGCAGCAGCAATTTCTGTGTGAAGAATCAAGCAGGAGAATTTG	600
Db	655	AAACCACTCCAAGGGCAGCAGCAATTTCTGTGTGAAGAATCAAGCAGGAGAATTTG	714
QY	601	GACGCGGAGATTTGAGATTTGACAGCAAGACATGTCTGCTCTGATTTACTCAGGAAC	660
Db	715	GACGCGGAGATTTGAGATTTGACAGCAAGACATGTCTGCTCTGATTTACTCAGGAAC	774
QY	661	GTCCTCAGGGGGAGAAAGCCCTTGGCTGGTCTAAATAGTGGGCTGTACACACATCACAG	720
Db	775	GTCCTCAGGGGGAGAAAGCCCTTGGCTGGTCTAAATAGTGGGCTGTACACACATCACAG	834
QY	721	CCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGGGGCTCAGTCCGCTGGTCTG	780
Db	835	CCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGGGGCTCAGTCCGCTGGTCTG	894
QY	781	CTTGTAACTCTACTCAACTCAGAATGAAGTAGCTGACAGCACTGGCTGAGGCTGGAGTTG	840
Db	895	CTTGTAACTCTACTCAACTCAGAATGAAGTAGCTGACAGCACTGGCTGAGGCTGGAGTTG	954
QY	841	CAGTGTTCGCTTGGAAGGGCGAGTCAGAAGATGACTTCTGGTGTATTTGACCCGCTGTG	900
Db	955	CAGTGTTCGCTTGGAAGGGCGAGTCAGAAGATGACTTCTGGTGTATTTGACCCGCTGTG	1014
QY	901	TGAACATGATGGGTGGCGAGGCCAACATGATCCTGGATGATGGGGAGACTTAAACCACT	960
Db	1015	TGAACATGATGGGTGGCGAGGCCAACATGATCCTGGATGATGGGGAGACTTAAACCACT	1074
QY	961	GGGTATTAAGAAATATCCAAACGTGTTAAGAAGATCCGAGGCAATTGTGAAGAGAGCG	1020
Db	1075	GGGTATTAAGAAATATCCAAACGTGTTAAGAAGATCCGAGGCAATTGTGAAGAGAGCG	1134
QY	1021	TGACTGTCTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTTCCGGCCA	1080
Db	1135	TGACTGTCTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTTCCGGCCA	1194
QY	1081	TGAACGTCAATGATTTCTGTACCACAGAAAGTTGATAACTGTACTGCTGCCGAGAAT	1140
Db	1195	TGAACGTCAATGATTTCTGTACCACAGAAAGTTGATAACTGTACTGCTGCCGAGAAT	1254
QY	1141	CCATTTTGGATGGCCTGAAGAGGACACAGATGTGATGTTGGTGGGAAACAAGTGTGG	1200
Db	1255	CCATTTTGGATGGCCTGAAGAGGACACAGATGTGATGTTGGTGGGAAACAAGTGTGG	1314
QY	1201	TGTGTGGCTATGGTGAAGGTAGGCAAGGGCTGTGTGCTGTCTCAAAAGCTCTTGAGCAA	1260
Db	1315	TGTGTGGCTATGGTGAAGGTAGGCAAGGGCTGTGTGCTGTCTCAAAAGCTCTTGAGCAA	1374
QY	1261	TGTCTACATTAACCGAAATCGAACCCCATCTGTGCTCTGACAGGCTGCATGGATGGTTCA	1320
Db	1375	TGTCTACATTAACCGAAATCGAACCCCATCTGTGCTCTGACAGGCTGCATGGATGGTTCA	1434
QY	1321	GGGTGCTAAAGCTAAATGAAGTCATCCGGCAAGTGCATGTGTAATAACTTGACAGGAA	1380
Db	1435	GGGTGCTAAAGCTAAATGAAGTCATCCGGCAAGTGCATGTGTAATAACTTGACAGGAA	1494
QY	1381	ATAAGATGTAGTGACACGGGAGCACTTGCATCCGATGAATAACAAGTTGATCGTATGCA	1440
Db	1495	ATAAGATGTAGTGACACGGGAGCACTTGCATCCGATGAATAACAAGTTGATCGTATGCA	1554
QY	1441	ATATGGGCCACTCCACACAGAAATCGATGTGACCAGGCTCCGCACTCCGAGCTGACGT	1500

Db	1555	ATATGGGCCACTCCACACAGAAATCGATGTGACCAGCCCTCCGCACTCCGGAGCTGACGT	1614
QY	1501	GGGAGCGAGTACGTTCTCAGGTGGACCATGTCACTCTGGCCAGATGGCAAACGAGTTGTCC	1560
Db	1615	GGGAGCGAGTACGTTCTCAGGTGGACCATGTCACTCTGGCCAGATGGCAAACGAGTTGTCC	1674
QY	1561	TCCTGGCAGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTTCCACCTTTGTCTGT	1620
Db	1675	TCCTGGCAGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTTCCACCTTTGTCTGT	1734
QY	1621	CCATCACAGCCACACACAGCGCTTTGGCAGCTGATAGAACTCTATAATGCAACCCGAGGGCC	1680
Db	1735	CCATCACAGCCACACACAGCGCTTTGGCAGCTGATAGAACTCTATAATGCAACCCGAGGGCC	1794
QY	1681	GATACAAGCAGAGTGTGTACTTGCTTCCTAAGAAATGGAATGAAATACGTGGCCAGCTTGC	1740
Db	1795	GATACAAGCAGAGTGTGTACTTGCTTCCTAAGAAATGGAATGAAATACGTGGCCAGCTTGC	1854
QY	1741	ATCTGCCATCATTTGATGCCCCACCTTACAGAGCTGACAGATGACCAAGCAAATATCTGG	1800
Db	1855	ATCTGCCATCATTTGATGCCCCACCTTACAGAGCTGACAGATGACCAAGCAAATATCTGG	1914
QY	1801	GACTCAACAAAAATGGGCCATTCAAACCTAATTATTACAGATACTAATGGACATACTAC	1860
Db	1915	GACTCAACAAAAATGGGCCATTCAAACCTAATTATTACAGATACTAATGGACATACTAC	1974
QY	1861	CAAGGACCACTCCACCTGAACACACACACTTAAGAATAATTTTAAAGATACTTTAT	1920
Db	1975	CAAGGACCACTCCACCTGAACACACACACTTAAGAATAATTTTAAAGATACTTTAT	2034
QY	1921	TTTCTTCTTACTCTCTTCTCTCTGATTTTTCCTATAATTTCATCTGTGTTTTCATC	1980
Db	2035	TTTCTTCTTACTCTCTTCTCTCTGATTTTTCCTATAATTTCATCTGTGTTTTCATC	2094
QY	1981	TCATTATCCAAGTTCTGCAGACCACACAGAACTTGCTTCATGGCTCTTGAATGAATA	2040
Db	2095	TCATTATCCAAGTTCTGCAGACCACACAGAACTTGCTTCATGGCTCTTGAATGAATA	2154
QY	2041	GAAGTTCAGGGTCCCTCACTCTAGTCACTAAGAAGATTTTACTCCCCAGCCAGAAA	2100
Db	2155	GAAGTTCAGGGTCCCTCACTCTAGTCACTAAGAAGATTTTACTCCCCAGCCAGAAA	2214
QY	2101	GGTGATCTCTCTCTTTTACCATTCTCTGGGACTTLAGCTTAAATTAGTACCTTATTACA	2160
Db	2215	GGTGATCTCTCTCTTTTACCATTCTCTGGGACTTLAGCTTAAATTAGTACCTTATTACA	2274
QY	2161	GGAATGCTAAGGTACCTCTCTGTGGAACAATCTGCAATGTCTAAATCGCCTTAAAAA	2220
Db	2275	GGAATGCTAAGGTACCTCTCTGTGGAACAATCTGCAATGTCTAAATCGCCTTAAAAA	2334
QY	2221	GCCCATTTCTTACGTCTGAAATCAGTGTCTTTCACACTCTTCAGAGAACGAGGATGGT	2280
Db	2335	GCCCATTTCTTACGTCTGAAATCAGTGTCTTTCACACTCTTCAGAGAACGAGGATGGT	2394
QY	2281	ACCTACCCGGCAGGTAGGTAGATGTGGGTGTGCATGTTAATTTCCCTTAGAAGTTCCA	2340
Db	2395	ACCTACCCGGCAGGTAGGTAGATGTGGGTGTGCATGTTAATTTCCCTTAGAAGTTCCA	2454
QY	2341	AGCCCTGTTTCTCGCTAAAGGTGTATGTCCAGTTCAGAGATGTGTATAATGAGCATGG	2400
Db	2455	AGCCCTGTTTCTCGCTAAAGGTGTATGTCCAGTTCAGAGATGTGTATAATGAGCATGG	2514
QY	2401	CTTGTTAAGATCAGGAGGCCCACTTGATTTATAGTATAGCCCTTCCCTCCACCA	2460
Db	2515	CTTGTTAAGATCAGGAGGCCCACTTGATTTATAGTATAGCCCTTCCCTCCACCA	2574
QY	2461	GACTTGCTCAATTTTCGAGTTTAACTAGACTACACTCTAATTTGAGTTAATTTGTCC	2520
Db	2575	GACTTGCTCAATTTTCGAGTTTAACTAGACTACACTCTAATTTGAGTTAATTTGTCC	2634
QY	2521	TCTAGATTATTTCTGTGTCCAAAAAATTTTAAAAAATTTTAAAAA 2563	

Db	1103	TGAAGCTCAATGATTTCTGTTACCAACAGAAAGTTGATAACTGTACTGCTGCCGAGAAT	1162
QY	1141	CCATTTTGGATGGCCTGAAGAGGACCACAGATGTGATGTTGGTGGAAACAAGTGTGG	1200
Db	1163	CCATTTTGGATGGCCTGAAGAGGACCACAGATGTGATGTTGGTGGAAACAAGTGTGG	1222
QY	1201	TGCTGTGCTATGCTGAGGTAGGCAAGGCTGCTGTGCTGCTCTCAAGCTCTTGAGCAA	1260
Db	1223	TGCTGTGCTATGCTGAGGTAGGCAAGGCTGCTGTGCTGCTCTCAAGCTCTTGAGCAA	1282
QY	1261	TTCGCTACATTTACCGAAATCGACCCCATCTGTGCTCTGCGAGGCTGCATGGATGGTCA	1320
Db	1283	TTCGCTACATTTACCGAAATCGACCCCATCTGTGCTCTGCGAGGCTGCATGGATGGTCA	1342
QY	1321	GGGTGGTAAAGCTAAATGAAGTCATCCGGCAAGTCGATGCTGTAATACTTGACAGGAA	1380
Db	1343	GGGTGGTAAAGCTAAATGAAGTCATCCGGCAAGTCGATGCTGTAATACTTGACAGGAA	1402
QY	1381	ATAAGAATGTAGTGACACGGGAGCAGCTGGATCGCATGAAACAGTGTATCGTATGCA	1440
Db	1403	ATAAGAATGTAGTGACACGGGAGCAGCTGGATCGCATGAAACAGTGTATCGTATGCA	1462
QY	1441	ATATGGGCCACTCCAAACACAGAAATCGATGTGACCAGCCTCCGCACTCCGGAGCTGACGT	1500
Db	1463	ATATGGGCCACTCCAAACACAGAAATCGATGTGACCAGCCTCCGCACTCCGGAGCTGACGT	1522
QY	1501	GGGAGCGAGTACGTTCTCAGGTGAGCAGTGTATCTGTGGCCAGATGGCAACGAGTGTCC	1560
Db	1523	GGGAGCGAGTACGTTCTCAGGTGAGCAGTGTATCTGTGGCCAGATGGCAACGAGTGTCC	1582
QY	1561	TCCTGGCAGAGGGTCTGCTACTCAATTTGAGCTGCTCCACAGTTCCACCTTTGTCTGT	1620
Db	1583	TCCTGGCAGAGGGTCTGCTACTCAATTTGAGCTGCTCCACAGTTCCACCTTTGTCTGT	1642
QY	1621	CCATCACAGCCACAACACAGGCTTTGGCAGTATAGAACTCTATAATGCAACCCGAGGGGC	1680
Db	1643	CCATCACAGCCACAACACAGGCTTTGGCAGTATAGAACTCTATAATGCAACCCGAGGGGC	1702
QY	1681	GATACAAGCAGGATGTGTACTTGTCTCTTAAGAAATGATGAATACGTTGCCAGCTTGC	1740
Db	1703	GATACAAGCAGGATGTGTACTTGTCTCTTAAGAAATGATGAATACGTTGCCAGCTTGC	1762
QY	1741	ATCTGCCATCATTTGATGCCCCACCTTACAGAGCTGACAGATGACCAACAAATATCTGG	1800
Db	1763	ATCTGCCATCATTTGATGCCCCACCTTACAGAGCTGACAGATGACCAACAAATATCTGG	1822
QY	1801	GACTCAACAATAATGGGCCATTCAAACCTAATTATTACAGATACTAATGACCATACTAC	1860
Db	1823	GACTCAACAATAATGGGCCATTCAAACCTAATTATTACAGATACTAATGACCATACTAC	1882
QY	1861	CAAGGACAGTCCACCTGAAACACACACACTTAAGAATAATTTTAAAGATACTTTAT	1920
Db	1883	CAAGGACAGTCCACCTGAAACACACACACTTAAGAATAATTTTAAAGATACTTTAT	1942
QY	1921	TTTTCTCTTACTCTCTTCTCTCTGATTTTTTTCCTATAATTTCAATCTTGTTTTTCATC	1980
Db	1943	TTTTCTCTTACTCTCTTCTCTCTGATTTTTTTCCTATAATTTCAATCTTGTTTTTCATC	2002
QY	1981	TCATTAATCCAAGTTCTGACAGACACACAGAACTGCTTCATGGCTCTTTAGATGAATA	2040
Db	2003	TCATTAATCCAAGTTCTGACAGACACACAGAACTGCTTCATGGCTCTTTAGATGAATA	2062
QY	2041	GAACTTCAGGGTCCCTCACTCTAGTCACTAAGAAGGATTTACTCCCCAGCCAGAGAAA	2100
Db	2063	GAACTTCAGGGTCCCTCACTCTAGTCACTAAGAAGGATTTACTCTCCAGCCAGAGAAA	2122
QY	2101	GGTGAATCTCTCTTTACCAATTTCTGGGAGCTTAGTCTTAATAGGACTTATTAACA	2160
Db	2123	GGTGAATCTCTCTTTACCAATTTCTGGGAGCTTAGTCTTAATAGGACTTATTAACA	2181
QY	2161	GGAATGCTAAGGTACCTTCTCTGTGAACAATCTCAATGTCTAAATGCGCTTAAAGA	2200

Db	2182	GGAAATGCTAAGGTACCTTCTCTGTGGACACATCTGCATGTCTAAATCGCCTTAAGA	2241
OY	2221	GCCCCATTCTTAGCTGCTGGAATCAGTGCCTTTCACCTTCTTCAGAGAAGCAGGATG	2280
Db	2242	GCCCCATTCTTAGCTGCTGGAATCAGTGCCTTTCACCTTCTTCAGAGAAGCAGGATG	2301
OY	2281	ACCTACCCGGCAGGTAGGTTAGATGTGGGTGGTGCATGTTAATTTCCCTTAGAAGTTCCA	2340
Db	2302	ACCTACCCGGCAGGTAGGTTAGATGTGGGTGGTGCATGTTAATTTCCCTTAGAAGTTCCA	2361
OY	2341	AGCCCTGTTCCCTGCGTAAAGGTGGTATGTCCAGTTCAGAGATGTGTATATGAGCATGG	2400
Db	2362	AGCCCTGTTCCCTGCGTAAAGGTGGTATGTCCAGTTCAGAGATGTGTATATGAGCATGG	2421
OY	2401	CTTGTTAAGATCAGGAGGCCCCACTTGGATTATAGTATAGCCCTTCCCTCCACTCCACCA	2460
Db	2422	CTTGTTAAGATCAGGAGGCCCCACTTGGATTATAGTATAGCCCTTCCCTCCACTCCACCA	2481
OY	2461	GACTTGCTCATTTTTCGAGTTTAACTAGACTACACTCTAATTGAGTTTAATTTGTCC	2520
Db	2482	GACTTGCTCATTTTTCGAGTTTAACTAGACTACACTCTA-TTGAGTTTAATTTGTCC	2540
OY	2521	TCTAGGATTATTTCTGTGTGCCAAAAAATTTTAAAAA	2563
Db	2541	TCTAGGATTATTTCTGTGTGCCAAAAAATTTTAAAAA	2583

RESULT 4	
LOCUS	BC016942
DEFINITION	BC016942 2552 bp mRNA linear PRI 09-NOV-2001
ACCESSION	Homo sapiens, S-adenosylhomocysteine-hydrolase-like 1, clone
VERSION	MGC:21453 IMAGE:3450568, mRNA, complete cds.
KEYWORDS	BC016942
SOURCE	BC016942.1 GI:16877386
ORGANISM	MGC.
	Homo sapiens.
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 2552)
AUTHORS	Strausberg, R.
TITLE	Direct Submission
JOURNAL	Submitted (05-NOV-2001) National Institutes of Health, Mammalian

REMARK	NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT	Contact: MGC help desk

cdNA Library Preparation: Life Technologies, Inc.
cdNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcde@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>
Series: IRAK Plate: 20 Row: a Column: 15
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

FEATURES	Location/Qualifiers
source	1. .2552

```

/organism="Homo sapiens"
/db_xref="LocusID:10768"
/db_xref="taxon:9606"
/clone="MGC:21453 IMAGE:3450568"
/tissue_type="Placenta, choriocarcinoma"
/clone_id="NH_MGC_10"

```

```
/lab_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
338.1840  
/codon_start=1  
/product="S-adenosylhomocysteine hydrolase-like 1"  
/protein_id="AAH16942.1"  
/db_xref="GI:16877387"  
/translation="MATVTAPKKQIOFADDMQETKPTKGRSLSRISQSDTS  
YSSASVTDSDDEVSPREKQNTSKSSNFCVKNIKQAEFRREIEIAEQMSALIS  
LRKRAOGEKPLAGKIVGCTHIACTAVLIETLICALGACRMSACNISTONEVAAL  
AEAGVAFAWKGESEDDFWICIDRCVNMMDGQANMILDDGDLTHWYKYPNVEKKI  
RGIVEESVTVHRLYQISKAGKLCVPAMNNDVTKOKFDNLYCCRESILDGLRTTD  
VMEGKQVNVCGYGEVKGCCALCALGAIVYTEIDPICALQACMDGFRVVKLNEVI  
ROYDVITCTGNKNVYTRHLDRLMKNSCIVCMGHSNTEIDVTSILRPELTWERVRSO  
VDHVIWPDGRVVLAEGRLLNLSCSTVPTFVLISITATQALALIELYNAPEGRYKOD  
VLLPKMDEVASLHPSFDAHLTELTDDAKYLGINKNGPRFPYRYR"  
BASE COUNT      644 a      600 c      671 g      637 t  
ORIGIN  
Query Match      98.5%; Score 2525.2; DB 9; Length 2552;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2549; Conservative 0; Mismatches 3; Indels 2; Gaps 2;  
QY      8 GCAGGTCGAGCTCGAGCTGCTCTGCTCTCTGTGTGGCCACCCTGCTGTCGGC 67  
Db      1 GCAGGTCGAGCTCGAGCTGCTCTGCTCTCTGTGTGGCCACCCTGCTGTCGGC 60  
QY      68 TGGCTTGGCTGCGGACAGACAGAGCGTGGCCACAGCAGCAGCAGCAGCAGCAGT 127  
Db      61 TGGCTTGGCTGCGGACAGACAGAGCGTGGCCACAGCAGCAGCAGCAGCAGT 120  
QY      128 CGAGCGAGGGGCGGCGAGAGGCTGGCGATCGCGTGTGGAGGGCGCGCGCGGCGAG 187  
Db      121 CGAGCGAGGGGCGGCGAGAGGCTGGCGATCGCGTGTGGAGGGCGCGCGCGGCGAG 180  
QY      188 CGGGCGGGCGCCAGAGGGGGAAGAGGGGGGGGGGGGTCAGCCGCTGGCGGGCGGG 247  
Db      181 CGGGCGGGCGCCAGAGGGGGAAGAGGGGGGGGGGGGTCAGCCGCTGGCGGGCGGG 240  
QY      248 CGGGGAATGTGATGCTGCTGACGCGATGCCGCTGCCCGGGGTCGGGGAGAGCTGAAGCA 307  
Db      241 CGGGGAATGTGATGCTGCTGACGCGATGCCGCTGCCCGGGGTCGGGGAGAGCTGAAGCA 300  
QY      308 GGGCAGAGGATCGAGGAGCGCCGAGAGTACTCTTCATGCGCCACCGTCACCAAGCGCC 367  
Db      301 GGGCAGAGGATCGAGGAGCGCCGAGAGTACTCTTCATGCGCCACCGTCACCAAGCGCC 360  
QY      368 CAAGAGCAAAATCCAGTTTGTGATGACATGCAAGAGTTCACCAATTCGCCACCAAAAC 427  
Db      361 CAAGAGCAAAATCCAGTTTGTGATGACATGCAAGAGTTCACCAATTCGCCACCAAAAC 420  
QY      428 TGGCCGAGATCTTGTCTCGCTGATCTCAAGTCTCTCCACTGACAGCTACAGTTGACG 487  
Db      421 TGGCCGAGATCTTGTCTCGCTGATCTCAAGTCTCTCCACTGACAGCTACAGTTGACG 480  
QY      488 TGCATCTACACAGATAGCTGTGATGAGTTTCTCCCCGAGAGAGAGCAAGCAAAACCA 547  
Db      481 TGCATCTACACAGATAGCTGTGATGAGTTTCTCCCCGAGAGAGAGCAAGCAAAACCA 540  
QY      548 CTCGAAGGGCAGCAGCAATTTCTGTGTGAAGAACATCAAGCAGGAGCAATTTGGACGGCG 607  
Db      541 CTCGAAGGGCAGCAGCAATTTCTGTGTGAAGAACATCAAGCAGGAGCAATTTGGACGGCG 600  
QY      608 GGAGATGAGATTCAGAGCAGACATGCTGCTGTGATTTCACTCAGGAAACGTGCTCA 667  
Db      601 GGAGATGAGATTCAGAGCAGACATGCTGCTGTGATTTCACTCAGGAAACGTGCTCA 660  
QY      668 GGGGAGAAAGCCCTGGCTGTGCTAAATAGTGGCTGTACACACATCAGCCAGAC 727  
Db      661 GGGGAGAAAGCCCTGGCTGTGCTAAATAGTGGCTGTACACACATCAGCCAGAC 720  
QY      728 AGCGGTGTGATGAGACACTGTGCTGCGGCTCAGTGCCTGTGCTGTAA 787  
Db      721 AGCGGTGTGATGAGACACTGTGCTGCGGCTCAGTGCCTGTGCTGTAA 780
```

```
780  
QY      788 CATCTACTCACTCAGAAATGAAGTAGCTGCAGCACCTGGCTGAGGCTGAGGTGCACTGT 847  
Db      781 CATCTACTCACTCAGAAATGAAGTAGCTGCAGCACCTGGCTGAGGCTGAGGTGCACTGT 840  
QY      848 CGCTTGAAGGGCGAGTCAGAAAGTAGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 907  
Db      841 CGCTTGAAGGGCGAGTCAGAAAGTAGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900  
QY      908 GGATGGGTGGCAGGCCAACATGATCTGATGATGGGGGAGACTTAACCCACTGGGTTTA 967  
Db      901 GGATGGGTGGCAGGCCAACATGATCTGATGATGGGGGAGACTTAACCCACTGGGTTTA 960  
QY      968 TAAGAGTATCCAAACGTGTTTAAGAAGATCCGAGGCATGTGGAAGAGAGCGTACTGG 1027  
Db      961 TAAGAGTATCCAAACGTGTTTAAGAAGATCCGAGGCATGTGGAAGAGAGCGTACTGG 1020  
QY      1028 TGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAGGCTGTGTGTGTGTGTGTGTGTGT 1087  
Db      1021 TGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAGGCTGTGTGTGTGTGTGTGTGTGT 1080  
QY      1088 CAATGATCTGTATACCAAGAGTTGATTAACCTGTACTGCTGCCGAGAAATCCATTTT 1147  
Db      1081 CAATGATCTGTATACCAAGAGTTGATTAACCTGTACTGCTGCCGAGAAATCCATTTT 1140  
QY      1148 GGATGGCTGAAGAGACACAGATGTGATGTTTGTGGGAAACAAGTGTGTGTGTGTGTGT 1207  
Db      1141 GGATGGCTGAAGAGACACAGATGTGATGTTTGTGGGAAACAAGTGTGTGTGTGTGTGT 1200  
QY      1208 CTATGTTAGGTAGGCAAGGCTGCTGTGCTCTCAAGCTCTTGGAGCAATGTGCTTA 1267  
Db      1201 CTATGTTAGGTAGGCAAGGCTGCTGTGCTCTCAAGCTCTTGGAGCAATGTGCTTA 1260  
QY      1268 CATTACCGAAATCGACCCCATCTGTGCTCTGACAGGCTGCTGATGGGTTCAGGGTGT 1327  
Db      1261 CATTACCGAAATCGACCCCATCTGTGCTCTGACAGGCTGCTGATGGGTTCAGGGTGT 1320  
QY      1328 AAAGCTAAATGAAGTCAATCCGCGCAAGTGCATGCTGTAATACTTGCACAGGAATAAGAA 1387  
Db      1321 AAAGCTAAATGAAGTCAATCCGCGCAAGTGCATGCTGTAATACTTGCACAGGAATAAGAA 1380  
QY      1388 TGTAGTGACACGGGAGCACTGGATGCGATGAAAAACAGTGTATGCAATATGGG 1447  
Db      1381 TGTAGTGACACGGGAGCACTGGATGCGATGAAAAACAGTGTATGCAATATGGG 1440  
QY      1448 CCACTCCACACAGAAATCGATGTGACCAAGCCTCCGACCTCCGAGCTGACGTGGAGCG 1507  
Db      1441 CCACTCCACACAGAAATCGATGTGACCAAGCCTCCGACCTCCGAGCTGACGTGGAGCG 1500  
QY      1508 AGTACGTTCTCAGGTGGACATGTCTATCTGGCCAGATGGCAACAGATTGCTCTCTGGC 1567  
Db      1501 AGTACGTTCTCAGGTGGACATGTCTATCTGGCCAGATGGCAACAGATTGCTCTCTGGC 1560  
QY      1568 AGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTTCACACCTTTGTTCTGTCCATCAC 1627  
Db      1561 AGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTTCACACCTTTGTTCTGTCCATCAC 1620  
QY      1628 AGCCACACACAGGCTTTGGCACTGATGAACTCTATAATGACCCGAGGGCGATACAA 1687  
Db      1621 AGCCACACACAGGCTTTGGCACTGATGAACTCTATAATGACCCGAGGGCGATACAA 1680  
QY      1688 GCAGATGTGTACTGCTTCCTAAGAAAAATGATGAATAGCTTGGCAGCTTGATCTGGC 1747  
Db      1681 GCAGATGTGTACTGCTTCCTAAGAAAAATGATGAATAGCTTGGCAGCTTGATCTGGC 1740  
QY      1748 ATCATTTGATGCCCACTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAA 1807  
Db      1741 ATCATTTGATGCCCACTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAA 1800  
QY      1808 CAAAAATGGGCATTAACCTTAATTAATACAGATACTAATGACCATACTACCAAGGAC 1867  
Db      1801 CAAAAATGGGCATTAACCTTAATTAATTAACAGATACTAATGACCATACTACCAAGGAC 1860
```

```

OY 1868 CAGTCCACCTGAACACACACTCTAAGAATATTTTAAGATACTTTTATTTCTTC 1927
    |||||||
DB 1861 CAGTCCACCTGAACACACACTCTAAGAATATTTTAAGATACTTTTATTTCTTC 1920
OY 1928 TTACTCTCTCTCTCTGATTTTCTCTATATTTTCATCTCTGTTTTCATCTCATAT 1987
    |||||||
DB 1921 TTACTCTCTCTCTCTGATTTTCTCTATATTTTCATCTCTGTTTTCATCTCATAT 1980
OY 1988 CCAAGTTCTGCAGACACACAGGAAGTCTGCTTCATGGCTCTTTAGATGAATAGAGTTC 2047
    |||||||
DB 1981 CCAAGTTCTGCAGACACACAGGAAGTCTGCTTCATGGCTCTTTAGATGAATAGAGTTC 2040
OY 2048 AGGTCCTCTACTCTAGTCACTAAGAAGGATTTTACTCCCGCAGCCGAGAAAGGTGAT 2107
    |||||||
DB 2041 AGGTCCTCTACTCTAGTCACTAAGAAGGATTTTACTCCCGCAGCCGAGAAAGGTGAT 2100
OY 2108 CTCTCTCTTACCATTTCTGGGAGCTTTAGTCTTAATTAAGTACCTTATTAACAGGAATG 2167
    |||||||
DB 2101 CTCTCTTACCATTTCTGGGAGCTTTAGTCTTAATTAAGTACCTTATTAACAGGAATG 2159
OY 2168 CTAAGTACCTCTCTCTGGAACAATCTGCAATGTCTAAATCGCCTTAAAGAGCCCAT 2227
    |||||||
DB 2160 CTAAGTACCTCTCTCTGGAACAATCTGCAATGTCTAAATCGCCTTAAAGAGCCCAT 2219
OY 2228 TCTTAGCTGCTGAATCAGTGTCTCTTCACTCTTCAAGAGAGGAGGATGTACCTACC 2287
    |||||||
DB 2220 TCTTAGCTGCTGAATCAGTGTCTCTTCACTCTTCAAGAGAGGAGGATGTACCTACC 2279
OY 2288 CGGAGGTAGGTAGATGTGGGTGTCATGTAAATTCCTTGAAGTCCAGGCCCTG 2347
    |||||||
DB 2280 CGGAGGTAGGTAGATGTGGGTGTCATGTAAATTCCTTGAAGTCCAGGCCCTG 2339
OY 2348 TTTCCTGCTAAAGGTGTATGTCTCAGTTCAGAGAGTGTAAATGAGCATGGCTGTGA 2407
    |||||||
DB 2340 TTTCCTGCTAAAGGTGTATGTCTCAGTTCAGAGAGTGTAAATGAGCATGGCTGTGA 2399
OY 2408 AGATCAGAGAGGCCCACTGGATTATAGTATAGCCCTCTCCACTCCACAGACTTGC 2467
    |||||||
DB 2400 AGATCAGAGAGGCCCACTGGATTATAGTATAGCCCTCTCCACTCCACAGACTTGC 2459
OY 2468 TCATTTTTCGAGTTTAACTAGACTACACTTATTTGAGTTTAAATTTTGTCTCTAGGA 2527
    |||||||
DB 2460 TCATTTTTCGAGTTTAACTAGACTACACTTATTTGAGTTTAAATTTTGTCTCTAGGA 2518
OY 2528 TTATTTCTGTGTCTCAAAAAA 2561
    |||||||
DB 2519 TTATTTCTGTGTCTCAAAAAA 2552
    |||||||

RESULT 5
BC007576 2526 bp mRNA linear PRI 12-JUL-2001
LOCUS Homo sapiens, S-adenosylhomocysteine hydrolase-like 1, clone
DEFINITION MGC:15558 IMAGE:3139729, mRNA, complete cds.
ACCESSION BC007576
VERSION BC007576.1 GI:14043176
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 2526)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (10-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT Contact: MGC help desk.
Email: cgabbs@email.nih.gov
Tissue Procurement: ATCC

```

```

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 22 Row: h Column: 21
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Similarity but not identity to protein.

FEATURES
    source
        1..2526
            /organism="Homo sapiens"
            /db_xref="locusID:10768"
            /db_xref="taxon:9606"
            /clone="MGC:15558 IMAGE:3139729"
            /tissue_type="Placenta, choriocarcinoma"
            /clone_id="NIH_MGC_21"
            /lab_host="DH10B-R"
            /note="vector: pOTB7"
            308..1810
                /codon_start=1
                /product="S-adenosylhomocysteine hydrolase-like 1"
                /protein_id="AAH07576.1"
                /db_xref="GI:14043177"
                /translation="MATVTKAPKKQIOFADMOEFTKPKTGRSLRSISOSTDS
                YSSAASYTDSDEVSPEKQOINSGSSNFCVKNIKQAEGRREIEAEDMSALIS
                LKRAQGEKPLAGAKIVGCTHITAQTAVLLETICALGACRWSACNIXSTONEVAAL
                AEAGVAFAMKGESEDDFWWCIDRCVNDGWOANMILDDGDLTHWYKKYPNVFKI
                RGIYESVTGVHRLYQSKRAGKLCVPAMNVNDSVTKQKFDNLVCCRESILDLKRTD
                VMFGKQVNVVCGYGEVGRKCAALKAIAIVITEIDPICALQACMDGFRVVKLNEVI
                ROVDVITCTGNKNVATREHLDRMKNSCIVCMGHSNTEIDVSLTPELTWERVRSQ
                VDHIWPDGKRVVLLAEGRLLNLSCTVPTFVLSITATQALALELYNAPEGRYKQD
                VYLLPKMDEVVASLHLSFDAHLELTDDQAKYLGINKNGPRPNRYR"

BASE COUNT 644 a 592 c 660 g 630 t
ORIGIN
OY 38 GTTCTCTTGTGGCCACCGTCTGCTGCGGCTGCTTGGGCTGCCGAACAGAGCGGTG 97
    |||||||
DB 1 GTTCTCTTGTGGCCGCGTCTGCTGCGGCTGCTTGGGCTGCCGAACAGAGCGGTG 60
OY 98 GGCCACAGCACTCAGAAAGCCGACGCTCGACGACGAGGGCGCCGACAGAGGTGGCGGA 157
    |||||||
DB 61 GGCCACAGCACTCAGAAAGCCGACGCTCGACGACGAGGGCGCCGACAGAGGTGGCGGA 120
OY 158 TCGCGTGTGGAGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 217
    |||||||
DB 121 TCGCGTGTGGAGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
OY 218 GGGCGGCGGCTCAGCCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 277
    |||||||
DB 181 GGGCGGCGGCTCAGCCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
OY 278 GCTGCCCGGGGTGCGGAGAGCTGAAGCAGGCGCAAGGAGATCGAGGACCGCGAGAAGTA 337
    |||||||
DB 241 GCTGCCCGGGGTGCGGAGAGCTGAAGCAGGCGCAAGGAGATCGAGGACCGCGAGAAGTA 300
OY 338 CTCCTTCATGGCCACCGCTCAGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 397
    |||||||

```

```

Query Match 97.4%; Score 2495.6; DB 9; Length 2526;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2520; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

```

Db 301 CTCCTTCATGGCCACCCTCACCAGGCGCCCAAGCAAAATCCAGTTGCTGATGACAT 360
QY 398 GCAGAGTTCACCAAAATTCCTCCACCAAACTGGCCGAAGATCTTGTCTGCTGATCTC 457
Db 361 GCAGAGTTCACCAAAATTCCTCCACCAAACTGGCCGAAGATCTTGTCTGCTGATCTC 420
QY 458 ACAGTCTCCACTGACAGCTACAGTTCAGTCTGATCTCTACACAGATAGCTGTGATGA 517
Db 421 ACAGTCTCCACTGACAGCTACAGTTCAGTCTGATCTCTACACAGATAGCTGTGATGA 480
QY 518 GGTTCCTCCCGAGAGAGACGACCAAACTCCAAAGGCGAGCAATTTCTGTGAA 577
Db 481 GGTTCCTCCCGAGAGAGAGACGACCAAACTCCAAAGGCGAGCAATTTCTGTGAA 540
QY 578 GAACATCAAGCAGGCAAGATTTGGACGCCGGAGATTGAGATTGCAAGCAAGACATGTC 637
Db 541 GAACATCAAGCAGGCAAGATTTGGACGCCGGAGATTGAGATTGCAAGCAAGACATGTC 600
QY 638 TGCTCTGATTTCACTCAGGAAACGCTGCTCAGGGGAGAAACCTTGCTGCTAAAT 697
Db 601 TGCTCTGATTTCACTCAGGAAACGCTGCTCAGGGGAGAAACCTTGCTGCTAAAT 660
QY 698 AGTGGCTGTACACACATCACAGCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCT 757
Db 661 AGTGGCTGTACACACATCACAGCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCT 720
QY 758 GGGGCTCAGTGCCTGCTGCTGCTGTTAATCTACTCACTCAAGATGAAGTAGCTGC 817
Db 721 GGGGCTCAGTGCCTGCTGCTGCTGTTAATCTACTCACTCAAGATGAAGTAGCTGC 780
QY 818 AGCACTGGCTGAGGCTGAGTTCAGTGTGCTTGAAGGGCGAGTCAGAAGATGACTT 877
Db 781 AGCACTGGCTGAGGCTGAGTTCAGTGTGCTTGAAGGGCGAGTCAGAAGATGACTT 840
QY 878 CTGCTGTGATGTAACCCCTGTGTGAACATGGATGGGCGAGGCCAACATGATCTGGA 937
Db 841 CTGCTGTGATGTAACCCCTGTGTGAACATGGATGGGCGAGGCCAACATGATCTGGA 900
QY 938 TGATGGGGGAGACTTAACCCACTGGGTTTATAGAAGTATCCAAAGCTGTTAAGAAGAT 997
Db 901 TGATGGGGGAGACTTAACCCACTGGGTTTATAGAAGTATCCAAAGCTGTTAAGAAGAT 960
QY 998 CCGAGGCAATGTGGAAGAGAGCGTGAAGTGTGTACAGGCTGTATCAGCTCTCCAAAGC 1057
Db 961 CCGAGGCAATGTGGAAGAGAGCGTGAAGTGTGTACAGGCTGTATCAGCTCTCCAAAGC 1020
QY 1058 TGGGAAGCTCTGCTGCTCCGGCCATGAACGTCAATGATCTGTTAACCAACAGAAGTTGA 1117
Db 1021 TGGGAAGCTCTGCTGCTCCGGCCATGAACGTCAATGATCTGTGTTAACCAACAGAAGTTGA 1080
QY 1118 TAACTTGTACTGTGCGGAGATTCATTTGGATGGCCTGAAGAGGACCACAGATGTGAT 1177
Db 1081 TAACTTGTACTGTGCGGAGATTCATTTGGATGGCCTGAAGAGGACCACAGATGTGAT 1140
QY 1178 GTTGTGTGGAACAAGTGC 1237
Db 1141 GTTGTGTGGAACAAGTGC 1200
QY 1238 TGCTCTCAAAAGCTTGGAGCAATTTGTACATTAACGAAATCGAACCCTCTGTGCTCT 1297
Db 1201 TGCTCTCAAAAGCTTGGAGCAATTTGTACATTAACGAAATCGAACCCTCTGTGCTCT 1260
QY 1298 GCAGGCTGTGATGAGTGGGTTCAGGGTGTAAAGCTAAATGAATCATCCGGCAAGTGA 1357
Db 1261 GCAGGCTGTGATGAGTGGGTTCAGGGTGTAAAGCTAAATGAATCATCCGGCAAGTGA 1320
QY 1358 TGTGTAATACTTGCACAGGAATTAAGATGTAGTGACACGGGAGCACTTGATCGCAT 1417
Db 1321 TGTGTAATACTTGCACAGGAATTAAGATGTAGTGACACGGGAGCACTTGATCGCAT 1380
QY 1418 GAAAAACAGTGTGATGCTATGCAATATGGGCACTCCAAACACAGAAATCGATGTGACAG 1477
Db 1381 GAAAAACAGTGTGATGCTATGCAATATGGGCACTCCAAACACAGAAATCGATGTGACAG 1440

QY 1478 CCTCCGACTCCGGAGCTGACGTGGAGCGAGTAGCTTCTCAGGTGAGGACCAATCATCTG 1537
Db 1441 CCTCCGACTCCGGAGCTGACGTGGAGCGAGTAGCTTCTCAGGTGAGGACCAATCATCTG 1500
QY 1538 GCCAGATGGCAACGAGTGTCTCTCTGGCAGAGGGTCTCTACTCAATTTGAGCTGCTC 1597
Db 1501 GCCAGATGGCAACGAGTGTCTCTCTGGCAGAGGGTCTCTACTCAATTTGAGCTGCTC 1560
QY 1598 CACAGTTCACACCTTTGCTCTCTCTCATCACAGCCACACACAGAGGCTTTGGCAGTGA 1657
Db 1561 CACAGTTCACACCTTTGCTCTCTCTCATCACAGCCACACACAGAGGCTTTGGCAGTGA 1620
QY 1658 ACTCTATATGCAACCCGAGGGCGATACAGCAGGATGTGTACTTGTCTCTTAAGAAAT 1717
Db 1621 ACTCTATATGCAACCCGAGGGCGATACAGCAGGATGTGTACTTGTCTCTTAAGAAAT 1680
QY 1718 GGATGAATACGTTGCCAGCTTGATCTGATCTGCAATCTGATGATGCCACCTTACAGAGCTGAC 1777
Db 1681 GGATGAATACGTTGCCAGCTTGATCTGATCTGCAATCTGATGATGCCACCTTACAGAGCTGAC 1740
QY 1778 AGATGAACCAAGCAAAATATCTGGAGCTCAACAAATATGGCCATTCAAACCTAATATTA 1837
Db 1741 AGATGAACCAAGCAAAATATCTGGAGCTCAACAAATATGGCCATTCAAACCTAATATTA 1800
QY 1838 CAGATACTAATGAGACCATACTACCAAGGACAGCTCCACCTGAACACACACTCTAAGAA 1897
Db 1801 CAGATACTAATGAGACCATACTACCAAGGACAGCTCCACCTGAACACACACTCTAAGAA 1860
QY 1898 ATATTTTAAAGATACTTTATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1957
Db 1861 ATATTTTAAAGATACTTTATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1920
QY 1958 AATTTCATCTTGTGTTTTCATCTCATTAACCAAGTTCTGACAGACACACAGGAAGTTGC 2017
Db 1921 AATTTCATCTTGTGTTTTCATCTCATTAACCAAGTTCTGACAGACACACAGGAAGTTGC 1980
QY 2018 TTCAATGCTCTTTAGATGAATAGAAGTTCAAGGTCCTCACTAGTCACTAAAGAAAG 2077
Db 1981 TTCAATGCTCTTTAGATGAATAGAAGTTCAAGGTCCTCACTAGTCACTAAAGAAAG 2040
QY 2078 ATTTTACTCCCCAGCCAGAAAGGTGATCTCTCTCTTACCATTCTGGGAGCTTACT 2137
Db 2041 ATTTTACTCTCCAGCCAGAAAGGTGATCTCTCTCTTACCATTCTGGGAGCTTACT 2099
QY 2138 CTTAATTAAGTACTTATTAACAGGAATGCTAAGTACTTCTCTGTGGAACAATCTGC 2197
Db 2100 CTTAATTAAGTACTTATTAACAGGAATGCTAAGTACTTCTCTGTGGAACAATCTGC 2159
QY 2198 AATGCTAAATCGCCTTAAAGAGCCCATTTCTAGCTGCTGAATCAGTGTCTCTTAC 2257
Db 2160 AATGCTAAATCGCCTTAAAGAGCCCATTTCTAGCTGCTGAATCAGTGTCTCTTAC 2219
QY 2258 TTCTTCAGAGAAGCAGGATGTACTACCCGGCAGGTAGTTAGATGTGGTGTGAT 2317
Db 2220 TTCTTCAGAGAAGCAGGATGTACTACCCGGCAGGTAGTTAGATGTGGTGTGAT 2279
QY 2318 GTTAATTTCCCTTAGAAGTTCCAAGCCCTGTTCTCTGCTGAAGGTGTATGTCCAGTTC 2377
Db 2280 GTTAATTTCCCTTAGAAGTTCCAAGCCCTGTTCTCTGCTGAAGGTGTATGTCCAGTTC 2339
QY 2378 AGAGATGTATATATGAGCATGGCTGTGAAGATCAGAGGCCCACTTGATTTATAGTA 2437
Db 2340 AGAGATGTATATATGAGCATGGCTGTGAAGATCAGAGGCCCACTTGATTTATAGTA 2399
QY 2438 TAGCCCTCTCCACTCCACAGACTTGTCTCATTTTGAAGTTTAACTAGACTACAC 2497
Db 2400 TAGCCCTCTCCACTCCACAGACTTGTCTCATTTTGAAGTTTAACTAGACTACAC 2459
QY 2498 TCTATTTGAGTTAATTTGCTCTAGATTTAATTTCTGTGTCCAAAAAATTTTAAAT 2557
Db 2460 TCTA-TTGAGTTAATTTTGTCTCTAGATTTAATTTCTGTGTCCAAAAAATTTTAAAT 2518

QY 2558 AAAAAA 2563
DB 2519 AAAAAA 2524

RESULT 6
HSU82761

LOCUS 2258 bp mRNA linear PRI 17-DEC-2001

DEFINITION Homo sapiens S-adenosyl homocysteine hydrolase homolog (XPVKona)

ACCESSION U82761

VERSION U82761.1 GI:2852124

KEYWORDS

SOURCE Homo sapiens.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2258)
Cleaver, J.E., Afzal, V., Feeney, L., McDowell, M., Sadlinski, W.,
Volpe, J.P.G., Busch, D.B., Coleman, D.M., Ziffer, D.W., Yu, Y.,
Nagasawa, H. and Little, J.B.
Increased ultraviolet sensitivity and chromosomal instability
related to P53 function in the xeroderma pigmentosum variant
Cancer Res. 59 (5), 1102-1108 (1999)

TITLE

JOURNAL MEDLINE
PUBMED
99168517
10070969

REFERENCE 2 (bases 1 to 2258)
Volpe, J.P.G., McDowell, M., Jostes, R.F., Afzal, V., Sadlinski, W.,
Trask, B.J., Legerski, R. and Cleaver, J.E.
Complementation of chromosomal instability in the xeroderma
pigmentosum variant by a gene on human chromosome 1 with homology
to S-adenosyl homocysteine hydrolase
Unpublished

JOURNAL
REFERENCE 3 (bases 1 to 2258)
Volpe, J.P.G., McDowell, M. and Cleaver, J.E.
Direct Submission
Submitted (19-DEC-1996) Dermatology, UCSF, 3rd and Parnassus, Box
0750, San Francisco, CA 94143, USA

TITLE

JOURNAL

FEATURES
source
1. 2258
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/cell_type="fibroblasts"
/tissue_type="skin"
1. 2258
/gene="XPVKona"
48. 1550
/note="xeroderma pigmentosum variant"
/codon_start=1
/product="S-adenosyl homocysteine hydrolase homolog"
/protein_id="AAC01960.1"
/db_xref="GI:2852125"
/translation="MATVTKAPKKOIOFADDMQEFTEPTKTGRSLRSISQSTDS
YSSAASYTDSDDVEVSPREKQNTSKSSNFCYKNIKQAEFGREIEIAEQDMSALIS
LKRAGEKPLAGAKIVGCTHITAGTAVLIEICALGACRWSACNISTONEVAAL
AEAGVAVFAWKGESEDDFWWCIDRCVNMDEQAMLLDDGDLTHWYKKYPNVFKKI
RGIVESVTGVHRLYOLSKAGKLCVPAMVNDSTKQKEDNLCCRESILDLKRTD
VMEGKQVNVCGYGEVGGKCCALKALGAIYITEIDPICALQACMDGFRVVKLNEVI
ROYDVITCTGNKNVTRHLDPMKNSCIYCNMGSNTEIDVTSRLTPELTWERVRSQ
VDHVIWPDGKRVLVLAEGRLNLSCTVPTFVLSITATTOALALIELYNAPEGRYKOD
VILLPKMDEVYASLHLPFSDAHLTELTDQAKYILGLNKNKPFKPNYYRY"

BASE COUNT 603 a 513 c 542 g 600 t

ORIGIN

Query Match 86.6%; Score 2218.6; DB 9; Length 2258;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2254; Conservative 0; Mismatches 4; Indels 3; Gaps 3;

QY 298- AGCTGAAGCAGGCCAAGAGATCGAGACGCCGAGAGTACTCTTCATGCGCACCGTCA 357
DB 1 AGCTGAAGCAGGCCAAGAGATCGAGACGCCGAGAGTACTCTTCATGCGCACCGTCA 60

QY 358 CCAAGGCGCCCAAGAACCAATCCAGTTGCTGATGACATGAGAGAGTTCACCAATTC 417
DB 61 CCAAGGCGCCCAAGAACCAATCCAGTTGCTGATGACATGAGAGAGTTCACCAATTC 120

QY 418 CCACCAAACTGGCGGAAGATCTTTGCTCGCTCGATCTCAGTCCACTGACAGCT 477
DB 121 CCACCAAACTGGCGGAAGATCTTTGCTCGCTCGATCTCAGTCCACTGACAGCT 180

QY 478 ACAGTTGAGTGTGATCTTACACAGATAGCTGTGATGATGAGGTTTCTCCCGAGAAGC 537
DB 181 ACAGTTGAGTGTGATCTTACACAGATAGCTGTGATGATGAGGTTTCTCCCGAGAAGC 240

QY 538 AGCAAACCACTCCCAAGGCGAGCAGCAATTTCTGTGAAGAACATCAAGCAGAGAT 597
DB 241 AGCAAACCACTCCCAAGGCGAGCAGCAATTTCTGTGAAGAACATCAAGCAGAGAT 300

QY 598 TTGGACGCGCGGAGATTTGAGATTGACAGAGCAGACATGTCTGCTCTGATTTCACTCAGGA 657
DB 301 TTGGACGCGCGGAGATTTGAGATTGACAGAGCAGACATGTCTGCTCTGATTTCACTCAGGA 360

QY 658 AACGTGCTCAGGGGAGAGAGCCCTTGCTGTGCTAAATAGTGGCTGTACACACATCA 717
DB 361 AACGTGCTCAGGGGAGAGAGCCCTTGCTGTGCTAAATAGTGGCTGTACACACATCA 420

QY 718 CAGCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGGGGCTCAGTGGCGCTGGT 777
DB 421 CAGCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGGGGCTCAGTGGCGCTGGT 480

QY 778 CTGCTGTACACTTACTTCACTCAGATGAGATGAGTGTGAGCTGACAGCTGAGGCTGAG 837
DB 481 CTGCTGTACACTTACTTCACTCAGATGAGATGAGTGTGAGCTGACAGCTGAGGCTGAG 540

QY 838 TTGCAGTGTTCGCTTGGAAAGGCGGAGTCAAGAGATGACTTGTGTGATTGACCGCT 897
DB 541 TTGCAGTGTTCGCTTGGAAAGGCGGAGTCAAGAGATGACTTGTGTGATTGACCGCT 600

QY 898 GTGTGAACATGATGGGTGGCAGGCCAACATGATCTGTGATGATGGGGAGACTTAACC 957
DB 601 GTGTGAACATGATGGGTGGCAGGCCAACATGATCTGTGATGATGGGGAGACTTAACC 660

QY 958 ACTGGTTTATAGAAGATATCCAAAGCTGTTTAAAGATCCAGAGCATTTGTGAAGAGA 1017
DB 661 ACTGGTTTATAGAAGATATCCAAAGCTGTTTAAAGATCCAGAGCATTTGTGAAGAGA 720

QY 1018 GCGTACTGTGTTTCAACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTCCGG 1077
DB 721 GCGTACTGTGTTTCAACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTCCGG 780

QY 1078 CCATGACGTCAATGATTTCTGTACCAACAGAGATTTGATTAATCTGTACTGCTGCCGAG 1137
DB 781 CCATGACGTCAATGATTTCTGTACCAACAGAGATTTGATTAATCTGTACTGCTGCCGAG 840

QY 1138 AATCCATTTTGGATGGCTGGAAGAGACACAGATGTGATGTTGGTGGAAACAAGTGG 1197
DB 841 AATCCATTTTGGATGGCTGGAAGAGACACAGATGTGATGTTGGTGGAAACAAGTGG 900

QY 1198 TGTGTGTGCTATGTGTAGGTAGGCAAGGGCTGTGTGCTGTCTCAAAAGCTCTTGAG 1257
DB 901 TGTGTGTGCTATGTGTAGGTAGGCAAGGGCTGTGTGCTGTCTCAAAAGCTCTTGAG 960

QY 1258 CAATGTCTACATTACCGAATGACCCCATCTGTGCTGTGACAGGCTGCATGATGGGT 1317
DB 961 CAATGTCTACATTACCGAATGACCCCATCTGTGCTGTGACAGGCTGCATGATGGGT 1020

QY 1318 TCAGGTGTGTAAGCTTAATGAGTCAATCCGCAAGTGCATGCTGTAATACTGCACAG 1377
DB 1021 TCAGGTGTGTAAGCTTAATGAGTCAATCCGCAAGTGCATGCTGTAATACTGCACAG 1080

QY 1378 GAAATTAAGATGTAGTGACACGGGAGCACTTGATGCGATGAATAACAGTGTATCGTAT 1437
DB 1081 GAAATTAAGATGTAGTGACACGGGAGCACTTGATGCGATGAATAACAGTGTATCGTAT 1140

QY 1438 GCAATATGGGCGCTCCACACAGAAATGATGTGACCAAGCTCCGCACTCCGAGACTGA 1497
DB 1141 GCAATATGGGCGCTCCACACAGAAATGATGTGACCAAGCTCCGCACTCCGAGACTGA 1200
QY 1498 CGTGGGAGCGAGTACGTTCTCAGGTGACCATGTCTATCTGGCCAGATGGCAACGAGTGG 1557
DB 1201 CGTGGGAGCGAGTACGTTCTCAGGTGACCATGTCTATCTGGCCAGATGGCAACGAGTGG 1260
QY 1558 TCCTCCTGGCAGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTTCACACCTTTGTTTC 1617
DB 1261 TCCTCCTGGCAGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTTCACACCTTTGTTTC 1320
QY 1618 TGTCCATCACAGCCACACACAGGCTTGGCACTGATAGAACTCTATATGACCCGAGG 1677
DB 1321 TGTCCATCACAGCCACACACAGGCTTGGCACTGATAGAACTCTATATGACCCGAGG 1380
QY 1678 GGGCATACAGCAGGATGTGTACTTGTCTTCTTAAGAAATGATGAATACGTTGCCAGCT 1737
DB 1381 GGGCATACAGCAGGATGTGTACTTGTCTTCTTAAGAAATGATGAATACGTTGCCAGCT 1440
QY 1738 TGCATCTGCCATCATTTGATGCCACCTTAACAGAGCTGACAGATGACCAAGCAAAATATC 1797
DB 1441 TGCATCTGCCATCATTTGATGCCACCTTAACAGAGCTGACAGATGACCAAGCAAAATATC 1500
QY 1798 TGGGACTCAACAAAATGGGCCATTCAAACCTAATTAATACAGATACCTAATGGACCATAC 1857
DB 1501 TGGGACTCAACAAAATGGGCCATTCAAACCTAATTAATACAGATACCTAATGGACCATAC 1560
QY 1858 TACCAAGAGCAGCTCCACCTGAACACACACACCTAAGAAATATTTTAAAGATAACTTT 1917
DB 1561 TACCAAGAGCAGCTCCACCTGAACACACACACCTAAGAAATATTTTAAAGATAACTTT 1620
QY 1918 TATTTCTCTTCTACTCTCTCTCTGATTTTTTCTCTATATTTTCTCTGTTTTC 1977
DB 1621 TATTTCTCTTCTACTCTCTCTCTGATTTTTTCTCTATATTTTCTCTGTTTTC 1680
QY 1978 ATCTCATTTATCCAAAGTCTGACAGACACAGAACTGCTCATGGCTCTTAGATGAA 2037
DB 1681 ATCTCATTTATCCAAAGTCTGACAGACACAGAACTGCTCATGGCTCTTAGATGAA 1740
QY 2038 ATAGAAGTTCAGGGTCCCTCCTACTAGTCACTAAGAAAGATTTTACTCCCCAGCCAG 2097
DB 1741 ATAGAAGTTCAGGGTCCCTCCTACTAGTCACTAAGAAAGATTTTACTCCCCAGCCAG 1800
QY 2098 AAAGTGATCTTCTCTTACCATTCTGGGAGCTTAGTCTTAATTAGTACCTTAATTA 2157
DB 1801 AAAGTGATCTTCTTACCATTCTGGGAGCTTAGTCTTAATTAGTACCTTAATTA 1859
QY 2158 ACAGAAATGCTAAGGTACCTTCTGTGGAACAATCTGCAATGTCTAATCGCCTTAA 2217
DB 1860 ACAGAAATGCTAAGGTACCTTCTGTGGAACAATCTGCAATGTCTAATCGCCTTAA 1919
QY 2218 AGAGCCCATTTCTTAGCTGCTGAATCAGTGTCTTCTCACTCTTCCAGAGAGCAGGAT 2277
DB 1920 AGAGCCCATTTCTTAGCTGCTGAATCAGTGTCTTCTCACTCTTCCAGAGAGCAGGAT 1979
QY 2278 GGTACTTACCCGGCAGGTAGGTAGATGTGGGTGTCATGTTAATTTCCCTTAGAAGT 2337
DB 1980 GGTACTTACCCGGCAGGTAGGTAGATGTGGGTGTCATGTTAATTTCCCTTAGAAGT 2039
QY 2338 CCAAGCCCTGTTTCTCTGCGTAAAGGTGTATGTCCAGTTCAGAGATGTGTATAATGACA 2397
DB 2040 CCAAGCCCTGTTTCTCTGCGTAAAGGTGTATGTCCAGTTCAGAGATGTGTATAATGACA 2098
QY 2398 TGGCTGTTAAGATCAGAGAGCCCACTTGATTTATAGTATAGCCCTTCCCTCAGCTCCA 2457
DB 2099 TGGCTGTTAAGATCAGAGAGCCCACTTGATTTATAGTATAGCCCTTCCCTCAGCTCCA 2158
QY 2458 CCAAGCTTGTCTATTTTTCGAGTTTAACTAGACTACACTCTATTTGAGTTTAAATTTG 2517
DB 2159 CCAAGCTTGTCTATTTTTCGAGTTTAACTAGACTACACTGTA-TTGAGTTTAAATTTG 2217
QY 2518 TCCTCTAGGATTTATTTCTGTGTCCAAAAA 2558

DB 2218 TCCTCTAGGATTTATTTCTGTGTCCAAAAA 2258

RESULT 7
HSM800298
LOCUS
DEFINITION Homo sapiens mRNA; cDNA DKFZp564A1523 (from clone DKFZp564A1523);
partial cds.
ACCESSION AL049954
VERSION AL049954.1 GI:4884203

KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE 1 (bases 1 to 2510)
JOURNAL Wambutt, R., Heubner, D., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
Submitted (15-MAY-1999) MIPS, Am Klopferspitz 18a, D-82152
Martinsried, GERMANY

COMMENT

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFZp564A1523) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://www.mips.biochem.mpg.de/proj/cDNA/>.

FEATURES

source

1..2510
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="1, 15.55 CR from D1S248"
/clone="DKFZp564A1523"
/tissue_type="brain"
/clone_1b="564 (synonym: hfbz2). Vector pAMP1; host X1-2blue; sites NotI + SalI"
/dev_stage="fetal"
1..1796
/gene="DKFZp564A1523"
<1..1796
/gene="DKFZp564A1523"
/note="S-adenosyl homocysteine hydrolase homolog XPVKONA splice variant"
/codon_start=3
/product="hypothetical protein"
/protein_id="CAB43223.1"
/db_xref="GI:4884204"
/translation="GLTVELDSSVCVYLILLRHGRGSRGKICAEISQPALPLPQETAPSGSESNLRNROGRRAESSONLSRLVSAQNGRSDRWLWQLWPEIDSLIEISCFIQFADDMQEFKFKPTGRSLRSISISQSTDSYSSAASYTDSDEYSPREKQGTNSRGSSNECVKNIKQAEFGREIEIAQDMSALISLRKRAQGEKPLAGAKIYGC THIRQTAVLIELTLCALGAQCQWRSACNISTONEVAALAEAGVAVFAWKGSEDEFW WCIDRCVYMDGWQAMILLDDGDLTHWYKKYPNVFKIRGIYESVTVGHRXLQLSK AGKLCVPMNVNDSVTKOKFDNLVCCRESILDLKRTTDMFGKQVAVCGYGEVKG CCAALKAIGAIIVYTEIDPICALQACMDGERVYKLVNVIROVDVITCTGKNVYTRE HLDNRNSCIVCMNGHSNTEIDVTSLRPELTWERVRSOVDHVIWPDGRVYLLAEGR LNLSCSTVPTFVLSITATQALALIELYNAPDEGRYKQDVYLLPKKMDRYVASLHLP S FDAHLIELDDQAKYILGLNKNPFRPNYRY"

gene
CDS
polya_site 2490
BASE COUNT 664 a 556 c 618 g 672 t
ORIGIN

Query Match 84.3%; Score 2161.2; DB 9; Length 2510;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2185; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 374 GCAATCCAGTTGCTGATGACATGCAGAGGTTACCAAAATTTCCCAACCAAACTGGCCG 433
DB 323 GCAATCCAGTTGCTGATGACATGCAGAGGTTACCAAAATTTCCCAACCAAACTGGCCG 382

QY 434 AAGATCTTGTCTCGCTCGATCTCACAGTCTCTCCACTGACAGCTACAGTTACGTGCATC 493
|||||
Db 383 AAGATCTTGTCTCGCTCGATCTCACAGTCTCTCCACTGACAGCTACAGTTACGTGCATC 442
QY 494 CTACACAGATAGCTCTGATGATGAGGTTTCTCCCGAGAGAGACAGCAACCAACTCCAA 553
|||||
Db 443 CTACACAGATAGCTCTGATGATGAGGTTTCTCCCGAGAGAGACAGCAACCAACTCCAA 502
QY 554 GGGCAGCAGCAATTTCTGTGTGAAGAACATCAAGCAGGAGATTTGGACGCCGGAGAT 613
|||||
Db 503 GGGCAGCAGCAATTTCTGTGTGAAGAACATCAAGCAGGAGATTTGGACGCCGGAGAT 562
QY 614 TGAGATTGCAGAGCAAGACATGTCTGCTCTGATTTCACCTCAGGAAACGTGCTCAGGGGA 673
|||||
Db 563 TGAGATTGCAGAGCAAGACATGTCTGCTCTGATTTCACCTCAGGAAACGTGCTCAGGGGA 622
QY 674 GAAGCCCTGGCTGGTGTAAATATAGTGGGTGTACACACATCAGAGCCAGACAGCGGT 733
|||||
Db 623 GAAGCCCTGGCTGGTGTAAATATAGTGGGTGTACACACATCAGAGCCAGACAGCGGT 682
QY 734 GTGATTGAGACACTCTGTGCCCCGTGGGGCTCAGTGCCTGCTGCTGTGTAACATCTA 793
|||||
Db 683 GTGATTGAGACACTCTGTGCCCCGTGGGGCTCAGTGCCTGCTGCTGTGTAACATCTA 742
QY 794 CTCAACTCAGAAATGAGTAGCTGCAGACACTGCGCTGAGGCTGAGTTGACAGTGTGCTG 853
|||||
Db 743 CTCAACTCAGAAATGAGTAGCTGCAGACACTGCGCTGAGGCTGAGTTGACAGTGTGCTG 802
QY 854 GAAGGCGAGTCAAGAAATGACTTCTGTTGTTGATTTGACCGCTGTGTGAACATGATGG 913
|||||
Db 803 GAAGGCGAGTCAAGAAATGACTTCTGTTGTTGATTTGACCGCTGTGTGAACATGATGG 862
QY 914 GTGCAGGCCCAACATGATCTGTGATGATGGGGAGACTTAAACCACCTGGGTTATAGAA 973
|||||
Db 863 GTGCAGGCCCAACATGATCTGTGATGATGGGGAGACTTAAACCACCTGGGTTATAGAA 922
QY 974 GTATCCAAACGTGTTTAAGAAATCCGAGGCAATTGTGAAGAGAGCGTGACTGTTTCA 1033
|||||
Db 923 GTATCCAAACGTGTTTAAGAAATCCGAGGCAATTGTGAAGAGAGCGTGACTGTTTCA 982
QY 1034 CAGGCTGATCAGCTCTCCAAAGCTGGGAGCTGTGTCTCCGGCCATGAACGTCAATGA 1093
|||||
Db 983 CAGGCTGATCAGCTCTCCAAAGCTGGGAGCTGTGTCTCCGGCCATGAACGTCAATGA 1042
QY 1094 TTCTGTTACCAACAGAAAGTTGATTAACCTGACTGCTGCCGAGAAATCCATTTGGATGG 1153
|||||
Db 1043 TTCTGTTACCAACAGAAAGTTGATTAACCTGACTGCTGCCGAGAAATCCATTTGGATGG 1102
QY 1154 CCTGAAGAGACACAGATGTGATGTTGGTGGGAAACAAGTGTGTGTGCTATGG 1213
|||||
Db 1103 CCTGAAGAGACACAGATGTGATGTTGGTGGGAAACAAGTGTGTGTGCTATGG 1162
QY 1214 TGAGTAGGCCAAGGGCTGCTGCTCTCAAAAGCTCTTGAGCAATTTGCTACATTAC 1273
|||||
Db 1163 TGAGTAGGCCAAGGGCTGCTGCTCTCAAAAGCTCTTGAGCAATTTGCTACATTAC 1222
QY 1274 CGAAATCGACCCCATCTGTGCTCTGCAGGCTGCATGGATGGTTCAAGGTTGTAAGCT 1333
|||||
Db 1223 CGAAATCGACCCCATCTGTGCTCTGCAGGCTGCATGGATGGTTCAAGGTTGTAAGCT 1282
QY 1334 AAATGAAGTATCCGGCAAGTGTGATGCTAATACTTGACAGAGAAATAGATAGT 1393
|||||
Db 1283 AAATGAAGTATCCGGCAAGTGTGATGCTAATACTTGACAGAGAAATAGATAGT 1342
QY 1394 GACACGGGAGCACTTGATCGCATGAAAAACAGTGTATCGTATGCAATATGGCCACTC 1453
|||||
Db 1343 GACACGGGAGCACTTGATCGCATGAAAAACAGTGTATCGTATGCAATATGGCCACTC 1402
QY 1454 CAACACAGAAATCGATGTGACCAAGCTCCGCACTCCGGAGCTGACGTGGGAGGAGTACG 1513
|||||
Db 1403 CAACACAGAAATCGATGTGACCAAGCTCCGCACTCCGGAGCTGACGTGGGAGGAGTACG 1462
QY 1514 TTCTCAGGTGACCATGTCTATCTGGCCAGATGGCAAAAGAGTTGTCTCTCTGGCAGAGGG 1573

|||||
Db 1463 TTCTCAGGTGACCATGTCTATCTGGCCAGATGGCAACAGAGTTGTCTCTCTGGCAGAGGG 1522
QY 1574 TCGTCTACTCAATTTGAGCTGCTCCACAGTGTCCACCTTTGTTCTGTCCATCACAGCCAC 1633
|||||
Db 1523 TCGTCTACTCAATTTGAGCTGCTCCACAGTGTCCACCTTTGTTCTGTCCATCACAGCCAC 1582
QY 1634 AACACAGGCTTTGGCACTGATAGAACTCTAATATGCAACCCGAGGGCGATACAGCAGA 1693
|||||
Db 1583 AACACAGGCTTTGGCACTGATAGAACTCTAATATGCAACCCGAGGGCGATACAGCAGA 1642
QY 1694 TGTGTACTGCTTCTTAAGAAATGATGAATACGTTGCCAGCTTGATCTGCCATCAAT 1753
|||||
Db 1643 TGTGTACTGCTTCTTAAGAAATGATGAATACGTTGCCAGCTTGATCTGCCATCAAT 1702
QY 1754 TGATGCCCACTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAACAAA 1813
|||||
Db 1703 TGATGCCCACTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAACAAA 1762
QY 1814 TGGGCCATCAACCTAATTATACAGATACTAATGACCATACTACCAAGGACAGTCC 1873
|||||
Db 1763 TGGGCCATCAACCTAATTATACAGATACTAATGACCATACTACCAAGGACAGTCC 1822
QY 1874 ACCTGAACCAACACTCTAAAGAAATATTTTAAATAACTTTATTTCTCTTACTC 1933
|||||
Db 1823 ACCTGAACCAACACTCTAAAGAAATATTTTAAATAACTTTATTTCTCTTACTC 1882
QY 1934 CTTCCTCTGATTTTCTCTATATTTCAATCTGTTGTTTTCATCTCATTTATCCAAGT 1993
|||||
Db 1883 CTTCCTCTGATTTTCTCTATATTTCAATCTGTTGTTTTCATCTCATTTATCCAAGT 1942
QY 1994 TCTGACAGCAACAGAACTTGTCTCATGCGCTTTAGATGAATAGAAATGATTCAGGCTC 2053
|||||
Db 1943 TCTGACAGCAACAGAACTTGTCTCATGCGCTTTAGATGAATAGAAATGATTCAGGCTC 2002
QY 2054 CCTCACTCTAGTCACTAAAGAAAGATTTTACTCCGCCAGCCAGAAAGGTGATTTCTCTC 2113
|||||
Db 2003 CCTCACTCTAGTCACTAAAGAAAGATTTTACTCTCCAGCCAGAAAGGTGATTTCT-TC 2061
QY 2114 TTTACCATTTCTGGGACTTTAGTCTTAATTAGTACCTTATTAACAGGAAATGCTAAG 2173
|||||
Db 2062 TTTACCATTTCTGGGACTTTAGTCTTAATTAGTACCTTATTAACAGGAAATGCTAAG 2121
QY 2174 TACCTTCTGTGGAACAATCTGCAATGTCTAAATGCGCTTAAAGAGCCCATTTCTAG 2233
|||||
Db 2122 TACCTTCTGTGGAACAATCTGCAATGTCTAAATGCGCTTAAAGAGCCCATTTCTAG 2181
QY 2234 CTGCTGAATCAGTGTCTTTCACTTCTTACAGAGCAGGAGTGTACCTACCCGGAG 2293
|||||
Db 2182 CTGCTGAATCAGTGTCTTTCACTTCTTACAGAGCAGGAGTGTACCTACCCGGAG 2241
QY 2294 GTAGTTAGATGTTGGTGGTGCATGTTAATTTCCCTTAGAAGTTCCAAGCCCTGTTCTC 2353
|||||
Db 2242 GTAGTTAGATGTTGGTGGTGCATGTTAATTTCCCTTAGAAGTTCCAAGCCCTGTTCTC 2301
QY 2354 GCGTAAAGTGTATGTCCAGTTCAGAGATGTGTATAATGAGCATGGCTTGTAAAGTCA 2413
|||||
Db 2302 GCGTAAAGTGTATGTCCAGTTCAGAGATGTGTATAATGAGCATGGCTTGTAAAGTCA 2361
QY 2414 GGAGGCCCATTTGATTTATAGTATAGCCCTCTCTCCACTCCACAGACTTGTCTAATT 2473
|||||
Db 2362 GGAGGCCCATTTGATTTATAGTATAGCCCTCTCTCCACTCCACAGACTTGTCTAATT 2421
QY 2474 TTGAGTTTAACTAGACTACACTATTTGAGTTAATTTTGTCTCTAGATTTATTT 2533
|||||
Db 2422 TTGAGTTTAACTAGACTACACTCTA-TTGAGTTAATTTTGTCTCTAGATTTATTT 2480
QY 2534 TCTGTGTCCAAAAA 2563
|||||
Db 2481 TCTGTGTCCAAAAA 2510
|||||

RESULT 8

BC018218		LOCUS	1772 bp	mRNA	linear	ROD 07-AUG-2002
		DEFINITION	Mus musculus, S-adenosylhomocysteine hydrolase-like 1, clone MGC:18748 IMAGE:4007102, mRNA, complete cds.			
		ACCESSION	BC018218			
		VERSION	BC018218.1	GI:17390492		
		KEYWORDS	MGC.			
		SOURCE	house mouse.			
		ORGANISM	Mus musculus			
		REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
		AUTHORS	1 (bases 1 to 1772)			
		TITLE	Strausberg, R.			
		JOURNAL	Direct Submission			
			Submitted (03-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA			
		REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov			
		COMMENT	Contact: MGC help desk Email: cgapbs-remail.nih.gov Tissue Procurement: Gilbert Smith, Ph.D. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/ Contact: amgdbcm.tmc.edu Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.			
		FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Series: IRAX Plate: 23 Row: n Column: 6 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Genomescan gene prediction, similarity but not identity to protein.			
		source	Location/Qualifiers 1..1772 /organism="Mus musculus" /db_xref="taxon:10090" /map="CZECB II" /clone="MGC:18748 IMAGE:4007102" /tissue_type="Mammary tumor metastatized to lung." MMTV-LTR/Mutl model. Expression driven by an MMTV-LTR enhancer." /clone_1lb="NCI_CGAP_Lu30" /lab_host="DH10B" /note="Vector: pCMV-SPORT6" 191..1693 /codon_start=1 /product="S-adenosylhomocysteine hydrolase-like 1" /protein_id="AAH18218.1" /db_xref="GI:17390493" /db_xref="LocustID:10768" /translation="MATVTAKPKKQIQPADMQEFTKEPTKTGRSLRSISQSSTDSYSSASYTSDDEVS PREKOOTNSKGSSNEFCVNIIKQAEFGREIEITAEQDMASALILKKRAQGEKPLAGAKIVGCTHITAQTAVLETICALGAOCWSACNISTONEVAALAEAGVAVEAFMKGESEDFFWMCIDRCVNMDCGMAMLLDDGDLTHWVKRYPNVEKKIRGIVESVTGVHRLYLQLSKAGKLCPANMYNDSVTKOKFDNLCCRESILDGLKRITDVMEGGKVAVCGYGEVGEGCCALKALAIYITEIDPICALOACMDGFERYKLINEVIROVDVITGTGNKNVYTRHEHLDRMKNSCIYCNGHSNTEIDVTSLRPELTWERVRSQVDHVITPDGRRVYLLAEGRLNLSCSTVPTEFYLSITATQAALIELVNAPEGRRYKQDYLLPKRMDERYASIHLPSEDAHLTELTDQAKYVLGLNKNGPFKPNYRY"			
CDS						
BASE COUNT	462 a	406 c	506 g	398 t		
ORIGIN						
Query Match	59.5%	Score 1525.4;	DB 10;	Length 1772;		
Best Local Similarity	92.7%;	Pred. No. 7e-299;				
Matches 1635; Conservative	0;	Mismatches 121;	Indels 7;	Gaps 3;		

QY	152	GGCGCATCCGCTGTCTCGGAGGGCGCCGCGGGGCGAGCGCGGGCGCCAGAGGGGGAAAG	211
Db	2	GGCGGAGCGCGTGGCCGAGGGCGCGCGCGG---GACGGGCGGGCGCCCGTGAGGGAAAG	57
QY	212	AGCGGGGGCGCGGGGTCAGCCGCTGGCCGGGGCGCGGGGGAATGTGATGCCCTGACGC	271
Db	58	AGCGGGGGCGGGGTTAGCCGCGGCGCGGGCGCGGGGATGTGATGCCCTGACGC	117
QY	272	GATGCCGCTGCCCGGGGTCGGGGAGGAGCTGAAGCAGCGCCAAAGAGATCGAGGACGCCGA	331
Db	118	GATGCCGCTGCCCGGGTGTCTGGGGAGGAGCTGAACACAGGCCAACAGAGATCGAGGACGCCGA	177
QY	332	GAAGTACTCCTTCATGGCCACCCTCACCAAGGCGCCCAAGAACCAATCCAGTTGCTGA	391
Db	178	GAAGTACTCCTTCATGGCCACCCTCACCAAGGCTCCCAAGAACCAATCCAGTTGCTGA	237
QY	392	TGACATGCAAGAGTTCACCAAAATTCGCCCAAAACTGGCCGAGATCTTGTCTCGCTC	451
Db	238	TGACATGCAAGAGTTCACCAAAATTCCTACTAAGACTGGCCGGAGATCTTGTCTCGTTC	297
QY	452	GATCTCACAGTCTCCACTGACAGCTACAGTTTCAAGTTCAGCTGCATCTCTACACAGATAGCTCTGA	511
Db	298	CATCTCACAAATCCTCCACAGACAGCTACAGTTTCAAGTTCAGCTGCATCTCTACAGATAGCTCTGA	357
QY	512	TGATGAGGTTCTCCCCGAGAGAAAGCAGCAAACTCCAAAGGCGCAGCAATTTCTG	571
Db	358	TGATGAGGTTCTCCCCGAGAGAAAGCAGCAAACTCGAAGGGCGCAGCAATTTCTG	417
QY	572	TGTGAAGAATCAAGCAGCGCAAAATTTGGACGCCGGAGATTGAGATTGCAGAGCAAGA	631
Db	418	TGTGAAGAATCAAGCAGCGCAAGATTGGACGCCGGAGATTGAGATTGCAGAGCAAGA	477
QY	632	CATGTCCTCTGATTTCACTCAGAAACGTGCTCAGGGGAGAAAGCCCTTGCTGTGTC	691
Db	478	CATGTCCTCTGATTTCACTCAGAAACGTGCTCAGGGAGAGAAAGCCCTTGCTGTGTC	537
QY	692	TAAATAGTGGGCTGTACACACATCACAGCCAGACAGCGGTGTGATTGAGACACTCTG	751
Db	538	TAAATAGTGGGCTGTACACACATCACAGCCAGACAGCGGTATTATGAGACCCCTTG	597
QY	752	TGCCCTGGGGGCTCAGTGGCGCTGCTCTGCTTGTACATCTCACTCAGAATGAAGT	811
Db	598	TGCCCTGGGAGCTCAGTGGCGCTGCTCTGCTGCAACATCTATTCACTCAGAATGAAGT	657
QY	812	AGCTGACGACCTGGCTGAGGCTGAGGTTGCAGTGTTCGCTTGAAGGGCGAGTCAGAGA	871
Db	658	AGCTGACGACCTGGCTGAGGCTGAGGTCGCGGTGTTGCTTGAAGGGCGAGTCAGAGA	717
QY	872	TGACTTCTGCTGTATTTGACCGCTGTGTGAACATGATGGGTGGCAGGCCAACATGAT	931
Db	718	TGATTTCTGCTGTGATTTGACCGCTGTGTCAACATGATGGGTGGCAGGCCAACATGAT	777
QY	932	CCTGGATGATGGGGAGACTTAACCCACTGGGTTTATAAGAGTATCCAAACGTGTTAA	991
Db	778	CCTGGATGATGGGGAGACTTAACCCACTGGGTTTATAAGAGTATCCAAACGTGTTAA	837
QY	992	GAAGATCCGAGGCATTTGGAAGAGAGCGGTGACTGTTCACAGGCTGTATCAGCTCTC	1051
Db	838	GAAGATCCGAGGCATTTGGAAGAGAGCGGTGACTGTTCACAGGCTGTATCAGCTCTC	897
QY	1052	CAAGCTGGGAAGCTGTGTTCGGGCCATGAACGTCAATGATTTCTGTACCAACAGAA	1111
Db	898	CAAGCTGGGAAGCTGTGTTCAGCCATGAATGTCAATGATTTCTGTACCAACAGAA	957
QY	1112	GTTTGATACTGTACTGCTGCCGAGAAATCCATTTTGATGGCTGAAGAGGACCAAGA	1171
Db	958	GTTTGATACTGTACTGCTGCCGAGAAATCCATTTTGATGGCCCTGAAGAGGACCAAGA	1017
QY	1172	TGTGATGTTTGGTGGGAAACAAGTGTGTGTGTGGCTATGTTGAGGTAGGCAAGGCTG	1231
Db	1018	TGTGATGTTTGGTGGGAAACAAGTGTGTGTGTGTGGCTATGTTGAGGTAGGCAAGGCTG	1077

QY 1232 CTGTGCTGCTCTCAAGCTCTTGAGCAATGTCTACATTAACGAAATCGACCCCATCTG 1291
|||||
Db 1078 CTGTGCTGCTCTCAAGCTCTTGAGCAATGTCTACATTAACGAAATCGACCCCATCTG 1137
QY 1292 TGTCTGACAGGCTGCTGATGGGTTCAGGGTGTAAAGCTAAATGAAGTCATCCGGCA 1351
|||||
Db 1138 TGTCTGACAGGCTGCTGATGGGTTCAGGGTGTAAAGCTAAATGAAGTCATCCGGCA 1197
QY 1352 AGTCGATGTCGTAATACTTGACAGGAATAAGATGTAGTGACACGGGAGCAGCTTGA 1411
|||||
Db 1198 GGTGACGCTGTGAATTAATGTCACAGGAATAAGATGTAGTGACCTGGGAGCAGCTTGA 1257
QY 1412 TCGCATGAATAACAGTGTATCGTATGCAATATGGCCACTCCACACAGAAATCGATGT 1471
|||||
Db 1258 CCGAATGAATAACAGTGTATCGTATGTCATATGGCCACTCCACACAGGAGATCGATGT 1317
QY 1472 GACCAAGCTCCGCACTCCGAGAGCTGACGTGGAGCGAGTACGTTCTCAGGTGACCATGT 1531
|||||
Db 1318 GACCAAGCTCCGCACTCCGAGAGCTGACGTGGAGCGGTACGTTCTCAGGTGACCATGT 1377
QY 1532 CATCTGGCCAGATGGCAAGAGAGTGTCTCTCTGCGAGAGGCTCTACTCAATTGAG 1591
|||||
Db 1378 CATCTGGCCAGATGGCAAGAGAGTGTCTCTCTGCGAGAGGCTCTACTCAATTGAG 1437
QY 1592 CTGCTCCACAGTCCCACTCTTGTCTCTGCTCATCACAGCCACACAGAGCTTTGGCACT 1651
|||||
Db 1438 CTGCTCCACAGTCCCACTCTTGTCTCTGCTCATCACAGCCACACAGAGCTTTGGCACT 1497
QY 1652 GATAGAACTCTATAATGCAAGCCGAGGGCGGATACAGAGAGTGTACTTGTCTCTAA 1711
|||||
Db 1498 GATAGAACTCTATAATGCAAGCCGAGGGCGGATACAGAGAGTGTACTTGTCTCTAA 1557
QY 1712 GAAATGATGAATACGTTGCCAGCTTGCATCTGCCATCATTTGATGCCACCTTACAGA 1771
|||||
Db 1558 GAAATGATGAATACGTTGCCAGCTTGCATCTGCCATCATTTGATGCCACCTTACAGA 1617
QY 1772 GCTGACAGATGACCAAGCAAAATATCTGGGACTCAACAAAATGGGCCATTCAAACTAA 1831
|||||
Db 1618 ACTGACAGATGACCAAGCAAAATATCTGGGACTCAACAAAATGGGCCATTCAAACTAA 1677
QY 1832 TTATTACAGATACTAATGACCATACTACCAAGGACAGTCCACTGAAACACAC-ACCTC 1890
|||||
Db 1678 TTATTACAGATACTAATGAC--ATAGTACAGTACAGCAGTCCACATGAACACGCAACTC 1735
QY 1891 TAAAGAAATATTTTAAAGATAA 1913
|||||
Db 1736 TAAAGAAATATTTTAAAGATAA 1758

RESULT 9
AF035319 1369 bp mRNA linear PRI 07-JUL-2000
LOCUS Homo sapiens clone 23931 mRNA, partial cds.
ACCESSION AF035319
VERSION AF035319.1 GI:2661082
KEYWORDS FLI-CDNA.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1369)
AUTHORS Andersson,B., Wentland,M.A., Ricafrente,J.Y., Liu,W. and Gibbs,R.A.
TITLE A 'double adaptor' method for improved shotgun library construction
JOURNAL Anal. Biochem. 236 (1), 107-113 (1996)
MEDLINE 96207227
PUBMED 8619474
REFERENCE 2 (bases 1 to 1369)
AUTHORS Yu,W., Andersson,B., Worley,K.C., Muzny,D.M., Ding,Y., Liu,W.,
Ricafrente,J.Y., Wentland,M.A., Lennon,G. and Gibbs,R.A.
TITLE Large-scale concatenation CDNA sequencing
JOURNAL Genome Res. 7 (4), 353-358 (1997)
MEDLINE 97264341
PUBMED 9110174

REFERENCE 3 (bases 1 to 1369)
AUTHORS Yu,W., Sarginson,J. and Gibbs,R.A.
TITLE Direct Submission
JOURNAL Submitted (20-NOV-1997) Molecular and Human Genetics, Baylor
College of Medicine, One Baylor Plaza S930, Houston, TX 77030, USA
FEATURES
source
1. 1369
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="23931"
/sex="female"
/tissue_type="brain"
/clone_lib="INIB"
/dev_stage="infant"
/note="This clone is similar to Drosophila M. Abd-B gene
encoded by GenBank Accession Number U07835. The I.M.A.G.E.
Consortium clone ID number is 23931"
CDS
1. 657
/note="similar to S-adenosylhomocysteine hydrolase"
/codon_start=1
/protein_id="AAB8189.1"
/db_xref="GI:2661083"
/translation="PRAMCGYGEVKGCCALKAIAIVITEIDPICALQACMDGFR
VKLINEVIRQYDVIVITCTGNKNVTRREHLDKMKNSICVCMGHSNTEIDVTSRTPFL
TWERVRSQVDHYIWDPEKRVVLAEGRLNLSCTVPEVLSITATQALIELYNA
PEGRYKQDVYLLPKMDEYVASLHLPSEDAHLTELTDDQAKYLGLNKNGPFKNYTRY
"

BASE COUNT 370 a 314 c 286 g 399 t
ORIGIN
Query Match 52.0%; Score 1333.8; DB 9; Length 1369;
Best Local Similarity 99.7%; Pred. No. 4.8e-260;
Matches 1357; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1201 TGTGTGCTATGCTGAGGTAGGCAAGGGCTGCTGCTGCTCTCAAAAGCTTTGAGCAA 1260
|||||
Db 11 TGTGTGCTATGCTGAGGTAGGCAAGGGCTGCTGCTGCTCTCAAAAGCTTTGAGCAA 70
QY 1261 TTGCTCACTTACCGAAATGCAACCCATCTGTGCTCTGACAGGCTCGATGGATGGTCA 1320
|||||
Db 71 TTGCTCACTTACCGAAATGCAACCCATCTGTGCTCTGACAGGCTCGATGGATGGTCA 130
QY 1321 GGGTGTAAAGCTAAATGAAGTCATCCGCAAGTGCATGCTGTAATTAAGTTCACAGAA 1380
|||||
Db 131 GGGTGTAAAGCTAAATGAAGTCATCCGCAAGTGCATGCTGTAATTAAGTTCACAGAA 190
QY 1381 ATAAGATGTAGTGACACGGGAGCACTTGATCGCATGAATAACAGTTGATCGTATGCA 1440
|||||
Db 191 ATAAGATGTAGTGACACGGGAGCACTTGATCGCATGAATAACAGTTGATCGTATGCA 250
QY 1441 ATATGGGCCACTCCACACAGAAATGATGTGACACAGCCCTCCGCACTCCGAGCTGACGT 1500
|||||
Db 251 ATATGGGCCACTCCACACAGAAATGATGTGACACAGCCCTCCGCACTCCGAGCTGACGT 310
QY 1501 GGGAGCGAGTACGTTCTCAGGTGACCATGTGCATCTGGCCAGATGGCAAGAGTGTCC 1560
|||||
Db 311 GGGAGCGAGTACGTTCTCAGGTGACCATGTGCATCTGGCCAGATGGCAAGAGTGTCC 370
QY 1561 TCCTGGCAGAGGGTCTCTACTCAATTGAGCTGTCTCCACAGTTCACACTTTGTCTGT 1620
|||||
Db 371 TCCTGGCAGAGGGTCTCTACTCAATTGAGCTGTCTCCACAGTTCACACTTTGTCTGT 430
QY 1621 CCATCACAGCCACACAGAGCTTTGGCACTGATAGAACTCTAATATGACCCGAGGGGC 1680
|||||
Db 431 CCATCACAGCCACACAGAGCTTTGGCACTGATAGAACTCTAATATGACCCGAGGGGC 490
QY 1681 GATACAGCAGAGTGTGACTGCTCTCTAAGAAATGATGATAGCTTGCAGCTTGC 1740
|||||
Db 491 GATACAGCAGAGTGTGACTGCTCTCTAAGAAATGATGATAGCTTGCAGCTTGC 550
QY 1741 ATCTGCCATCATTTGATGGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1800
|||||
Db 551 ATCTGCCATCATTTGATGGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 610

```

QY 1801 GACTCAACAAAAATGGCCATTCAAAACCTAATTATACAGATACTAATGACCATTACTAC 1860
    |||||||
Db 611 GACTCAACAAAAATGGCCATTCAAAACCTAATTATACAGATACTAATGACCATTACTAC 670
QY 1861 CAAGACCACTCCACCTGACACACACACACACACACACACACACACACACACACACAC 1920
    |||||||
Db 671 CAAGACCACTCCACCTGACACACACACACACACACACACACACACACACACACACAC 730
QY 1921 TTCTCTCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1980
    |||||||
Db 731 TTCTCTCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 790
QY 1981 TCATTAATCCAAAGTCTCTGACAGACACACAGAACTCTCATGGCTCTTAGATGAATA 2040
    |||||||
Db 791 TCATTAATCCAAAGTCTCTGACAGACACACAGAACTCTCATGGCTCTTAGATGAATA 850
QY 2041 GAAGTTCAGGGTCCCTCACTCTAGTCACTAAAGAGATTTTACTCCCCAGCCAGAAA 2100
    |||||||
Db 851 GAAGTTCAGGGTCCCTCACTCTAGTCACTAAAGAGATTTTACTCTCCAGCCAGAAA 910
QY 2101 GGTGATCTCTCTCTTACCATTCTCTGGGACTTACTCTTAATTAGGTACCTATTACA 2160
    |||||||
Db 911 GGTGATCTCTCTCTTACCATTCTCTGGGACTTACTCTTAATTAGGTACCTATTACA 969
QY 2161 GGAATATGCTAAGGTACCTCTCTGTGGAACAATCTGCAATGCTAAATCGCCCTTAAAGA 2220
    |||||||
Db 970 GGAATATGCTAAGGTACCTCTCTGTGGAACAATCTGCAATGCTAAATCGCCCTTAAAGA 1029
QY 2221 GCCCATTTCTAGCTCTGCTGAATCAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2280
    |||||||
Db 1030 GCCCATTTCTAGCTCTGCTGAATCAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1089
QY 2281 ACCTACCCGGCAGGTAGGTAGATGTGGGTGGTGCATGTTAATTCCCTTAGAAGTTCCA 2340
    |||||||
Db 1090 ACCTACCCGGCAGGTAGGTAGATGTGGGTGGTGCATGTTAATTCCCTTAGAAGTTCCA 1149
QY 2341 AGCCCTGTTCTCCGCTAAGGTGATGTCTCAGTTCAGAGATGTATATGAGCATGG 2400
    |||||||
Db 1150 AGCCCTGTTCTCCGCTAAGGTGATGTCTCAGTTCAGAGATGTATATGAGCATGG 1209
QY 2401 CTGTGTAAGATCAGGAGGCCCACTGGATTATATAGATAGCCCTCTCCACTCCACCA 2460
    |||||||
Db 1210 CTGTGTAAGATCAGGAGGCCCACTGGATTATATAGATAGCCCTCTCCACTCCACCA 1269
QY 2461 GACTTGCTCATTTTTCGAGTTTAACTAGACTACACTATTTGAGTTAATTTGTCC 2520
    |||||||
Db 1270 GACTTGCTCATTTTTCGAGTTTAACTAGACTACACTCTTA-TTGAGTTAATTTGTCC 1328
QY 2521 TCTAGAGTTATTTCTGTGTCTCCAAAAA 2561
    |||||||
Db 1329 TCTAGAGTTATTTCTGTGTCTCCAAAAA 1369

```

```

RESULT 10 1323 bp mRNA linear PRI 12-JUL-2001
LOCUS BC003631
DEFINITION Homo sapiens, similar to S-adenosylhomocysteine hydrolase-like 1,
clone IMAGE:3010755, mRNA, partial cds.
ACCESSION BC003631
VERSION BC003631.1 GI:13177702
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 1323)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (26-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

```

REMARK

COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 contact: amadansystemsbiology.org
 Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
 Greene, Mark Kettelman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAL Plate: 3 Row: b Column: 6.
 Location/Qualifiers

FEATURES

source

1. 1323
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3010755"
 /tissue_type="Muscle, rhabdomyosarcoma"
 /clone_1db="NIH_MGC_17"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 <1..611
 /codon_start=3
 /product="Similar to S-adenosylhomocysteine hydrolase-like
 1"
 /protein_id="AAH03631.1"
 /db_xref="gi:13177703"
 /translation="HEBALGAIIVITEIDPICALQACMDGFRVKNLEVIROYDVIT
 CTGNKRVYTRHLDKRMKNSICVNMGSHNEIDVTSIRPELTWERVRSQVDHIMPD
 GKRVLIAEGRLLNLSCSTVPTFVLSIRATQALALIELYNAPEGRYKQDVIYLLPKKM
 DEYVASLHPSFDAHLELTDDQAKYLGINKNPFKNYRY"

CDS

BASE COUNT 365 a 303 c 271 g 384 t
 ORIGIN

Query Match 50.3%; Score 1288.8; DB 9; Length 1323;
 Best Local Similarity 99.7%; Pred. No. 6.4e-251;
 Matches 1312; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

```

QY 1246 AAGCTCTTGAGCAATGTCTACATTACCGAAATCGACCCCATCTGTGCTGACAGGCT 1305
    |||||||
Db 10 AAGCTCTTGAGCAATGTCTACATTACCGAAATCGACCCCATCTGTGCTGACAGGCT 69
QY 1306 GCATGGATGGGTTACGGGTGTAAGCTAAATGAAGTATCCGGCAAGTCGATGCTGTA 1365
    |||||||
Db 70 GCATGGATGGGTTACGGGTGTAAGCTAAATGAAGTATCCGGCAAGTCGATGCTGTA 129
QY 1366 TAACCTGCACAGAAATAGAATGTAGTACACAGGAGCACTTGATCGCATGAAAAACA 1425
    |||||||
Db 130 TAACCTGCACAGAAATAGAATGTAGTACACAGGAGCACTTGATCGCATGAAAAACA 189
QY 1426 GTTGATCGTATGCAATATGGGCCACTCCACACACAGAAATCGATGTGACCAAGCCTCCGA 1485
    |||||||
Db 190 GTTGATCGTATGCAATATGGGCCACTCCACACACAGAAATCGATGTGACCAAGCCTCCGA 249
QY 1486 CTCGGAGCTGACGTGGGAGCGAGTACGTTCTCAGGTGGACCATGTCATCTGGCCAGATG 1545
    |||||||
Db 250 CTCGGAGCTGACGTGGGAGCGAGTACGTTCTCAGGTGGACCATGTCATCTGGCCAGATG 309
QY 1546 GCAACGAGTTGCTCTCTGCGCAGAGAGGCTCTACTCAATTTGAGCTGCTCCACAGTTC 1605
    |||||||
Db 310 GCAACGAGTTGCTCTCTGCGCAGAGAGGCTCTACTCAATTTGAGCTGCTCCACAGTTC 369
QY 1606 CCACCTTTGTTCTGTCATCACAGCCACACACAGAGCTTTGGCACTGATAGAACTCTATA 1665
    |||||||
Db 370 CCACCTTTGTTCTGTCATCACAGCCACACACAGAGCTTTGGCACTGATAGAACTCTATA 429
QY 1666 ATGCACCCGAGGGCGGATACAAGCAGAGATGTGTACTTGCCTTCTTAAGAAATGATGAAT 1725
    |||||||
Db 430 ATGCACCCGAGGGCGGATACAAGCAGAGATGTGTACTTGCCTTCTTAAGAAATGATGAAT 489

```


DB	QY	REFERENCE	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES	BASE COUNT
Db	1466	CGATGTGACCAAGCCTTCGGCACTCCGGAGCTGACGTGGGAGCGAGTACGTTCTCAGGTGGA	AK025372	Homo sapiens cDNA: FLJ21719 f1s, clone COLF0094.	AK025372	AK025372.1	GI:10437875	oligo capping; f1s (full insert sequence).	Homo sapiens	1987 bp	mRNA	1linear	PRI 29-SEP-2000					557 a 430 c 501 g 499 t
Db	1479	TGACGTGGCGAGTCTGCGGACACACAGAACTGACTGGGAGCGAGTGAATCTCAAGTTGA																
Qy	1526	CCATGTGATCTGGCCAGATGGCAACGAGTTGCTCTCTGCGAGAGGGTCTGCTACTCAA																
Db	1539	CCATGTGATATGGCTGATGGCAAGAGATAGTACTGCTGGCAGAGGGCCGCTGTA																
Qy	1586	TTTGAGCTGCTCCAGTCTCCACCTTTGTTCTGTCATCACAGCCACACAGAGCTTT																
Db	1599	CCTTAGCTGCTCCACAGTGCCTACATTTGTGCTCTCAATCACTGCTACTACTAGGCTCT																
Qy	1646	GGCACTGATAGAATCTATATATGCAACCCGAGGGGCGATACAGCAGATGTGTACTTGT																
Db	1659	TGCTTGTATAGAGCTTACATATGCTCTGAGGGTCTGCTATTAAGCAGGATGTCTACTGTT																
Qy	1706	TCCTAAGAAATGGATGATACGTTGCCAGCTTGCATCTGCCATCATTTGATGCCCACT																
Db	1719	GCCCAAGAGATGGATGATGTGTGGCCAGCCCTACACCTGCTTACCTTTGATGCCCACTT																
Qy	1766	TACAGAGCTGACAGATGACCAAGCAAAATATCTGGACTCAACAAAATGGGCCATCAA																
Db	1779	GACAGAGCTGACAGATGACAGAGCCCAAGTATCTGGACTCAACAGAATGGGCCCTTCAA																
Qy	1826	ACCTAATTATACATTAATGACCATACCTACTACCAAGGACCACT 1871																
Db	1839	GCCTAATTACTACAGGATTTAAGTTCCTGTAACCTCAACAGCAATT 1884																
RESULT 14																		
AK025372 1987 bp mRNA 1linear PRI 29-SEP-2000																		
DEFINITION Homo sapiens cDNA: FLJ21719 f1s, clone COLF0094.																		
ACCESSION AK025372																		
VERSION AK025372.1 GI:10437875																		
KEYWORDS oligo capping; f1s (full insert sequence).																		
SOURCE Homo sapiens colon mucosa cDNA to mRNA, clone_11b:COLF clone:COLF0094.																		
ORGANISM Homo sapiens																		
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.																		
AUTHORS Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.																		
TITLE NEDO human cDNA sequencing project																		
JOURNAL Unpublished																		
REFERENCE 2 (bases 1 to 1987)																		
AUTHORS Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shibahara,T., Tanaka,T. and Nakamura,Y.																		
TITLE Direct Submission																		
JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@leims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)																		
COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry for Biotechnology; cDNA library sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).																		
FEATURES																		
source																		
1.1987																		
/organism="Homo sapiens"																		
/db_xref="taxon:9606"																		
/clone="COLF0094"																		
/tissue_type="colon mucosa"																		
/clone_11b="COLF"																		
/note="cloning vector pME18SFL3"																		

Query Match	38.3%;	Score 982.2;	DB 9;	Length 1987;
Best Local Similarity	78.2%;	Pred. No. 9.6e-189;		
Matches 11/9;	Conservative	0;	Mismatches 328;	Indels 0;
			Gaps	0;
365	GCCCAAGAAGCAAAATCCAGTTTGGCTGATGACATGACAGAGTTTCAACAAATTTCCACCACAA	424		
316	GCTTGAGTGGATGATCCAGTTTGGCTGACCAAGAGCAAGAAATTCACAAACGTTCCACCACAA	375		
425	AACGGCCGAAGATCTTTGTCTCGCTCGATCTCACAGTCCCTCCACTGACAGCTACAGTTC	484		
376	AATTGACGTCGCTCTTGTCTCGTTCCATTCTCAGTCATCTACTGACAGCTACAGCTC	435		
485	AGCTGCATCTCACACAGATAGCTCTGATGATGAGGTTCTTCCCCGACAGAAAGCAGCAAC	544		
436	AGCGGCTCATATACAGATAGCTCTGATGATGAGACATCGCCACAGGACAAAGCAGCAACAA	495		
545	CAACTCCAAAGGCGACAGCAATTTCTGTGAGAAACATCAAGCAGCAGAAATTTGGACG	604		
496	GAACTTAAGGGAAGCAGTACTTCTGTGTTAAGAACATCAACACAGCAGAGTTTGGACG	555		
605	CCGGAGATTGAGATTGACAGACAGACATGTCCTCTGATTTCACTCAGGAAACGTC	664		
556	AAGAGAAATTGAATTTGCTGAACAGAAATGCTGCAATGATGGCTTTGAGGAAGAGAGC	615		
665	TCAGGGGAGAAAGCCCTTGGCTGGCTAAATAGTGGGCTGTACACATCACAGCCCA	724		
616	TCAGAGAGAAAGCCTTTGGCTGGAGCCAAATCGTGGTTGACACACATCACTGCTCA	675		
725	GACAGCGGTGTTGATTGAGACACTCTGTGCCCTGGGGCTCAGTGGCGTGTCTGCTG	784		
676	GACTGCTGTCTTATGGAACCTCTGGCTCTGGGGCCAGTGGCATGGGCTGCTG	735		
785	TAACATCTACTCACTCAGATGAGATGACTGACAGCACTGGCTGAGGCTGGAGTTGCAGT	844		
736	CAACATCTATTCACACTCTCAATGAGAGTGGCTGCTCTAGCAGAAAGTGATTTCTCTGT	795		
845	GTTGCTTGGAGGGGCGAGTCAAGATGACTTCTGGTGTATGACCCGCTGTGAA	904		
796	TTTTGCTGGAAAGGAGAGTCAAGATGACTTCTGGTGTATGATGATGATGTGAA	855		
905	CATGATGGGTGGAGGCCCAACATGATCTGATGATGGGGAGACTTAACCCACTGGGT	964		
856	TGTGAGGGCTGGCAGCCAAACATGATCTTGGATGATGAGGGGATTAACCCACTGGAT	915		
965	TTATTAAGAAATATCCAAACGTTTAAAGAAATCCGAGCATTTGGAAGAGAGCGTGAC	1024		
916	TTATTAAGAAATATCCCAACATGTTTAAAGAAATCAAGGCATAGTAGAGAGAGTGTAC	975		
1025	TGTGTTTACACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTGTCCGGCCATGAA	1084		
976	TGGAGTTTACACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTGTCCAGCCATGAA	1035		
1085	CGTCAATGATTTCTGTATCCAAACAGAAATTTGATACTTGTACTGTGCGGAGAAATCCAT	1144		
1036	TGTCAATGACTCAGTCACCAACAGAAATTTGACAACCTTACTGTGCGGTGAATCAAT	1095		
1145	TTTGGATGGCTGAAAGAGGACCAAGATGTGATGTTTGGTGGGAACAAGTGTGTGTG	1204		
1096	TCTTGATGAGACTTAAAGAGCAACAGACATGATGTTTGGTGGAAACAAAGTGTGTGTG	1155		
1205	TGGCTATGCTGAGTAGGCAAGGCTGTGTGCTGCTCTCAAGCTCTTGGAGCAATTTGT	1264		
1156	TGGCTATGAGAGAGGTGGGGAAGGGTCTGTGTGCTGCGCTGAAAGCCATGGGCTCCATTTGT	1215		
1265	CTACATTAACGAAATCGACCCCATCTGTCTGTGACAGGCTGCATGATGGGTTCAAGGT	1324		
1216	GTATGTAATGAAATGACCCCATCTGTGTGCGCTGCAAGCCTGTATGATGATGATTTGCACT	1275		
1325	GGTAAAGCTTAATGAAATGATCAATCCGCAAGTGTGATGTTGTAATTAATTTGCACAGGAAATTA	1384		
1276	GGTGAATTAATGAGGTCTATCCGCAAGTGTGACATTTGTATTAATCTGTACAGGTAAACAA	1335		

OY	1385	GAATGTAGTGACACGGGAGCAGCACTTGGATCGCATGAAAAAACAGTTGTATCGTATGCATAT	1444
Db	1336	GAAATGTGTTAACCAGAGAGAGCACTTGGACCGTATGAAGAATAAGCTGCATCGTTTGTAAACAT	1395
OY	1445	GGGCCACTCCAAACACAGAAAATCGATGTGACACGAGCCTCCGCACTCCGAGAGCTGACGTGGGA	1504
Db	1396	GGGACATTCCAACACACAGAGATTGACGTGGCGAGTCTGCGGACACCAAGAACTGACCTGGGA	1455
OY	1505	GCGAGTAGCTTCTCAGGTGAGACCATGTTCATCTGGCCAGATGGCAAACGAGTGTCTCCTCT	1564
Db	1456	GCGAGTAGATCTCAAGTTGACCATGTGATATGCGCTGATGGCAAGAGGATAGTACTGCT	1515
OY	1565	GCGAGAGGGTCTCTACTCAATTGAGCTGCTCCACAGTCCCACTTTGTTCTGTCCAT	1624
Db	1516	GCGAGAGGGCGCCCTGCTGTAACCTTAGCTGCTCCACAGTGCCTCAATTTGTGCTCTCAAT	1575
OY	1625	CACAGCCACAACACAGGCTTTGGCACTGATAGAACTCTATAATGCACCCGAGGGCGATA	1684
Db	1576	CACGTCTACTACTCAGGCTCTTGCCCTTGATAGAGCTTTACAATGCTCCTGAGGGTCGCTA	1635
OY	1685	CAAGCAGGATGTGTACTTGCCTTCCCTAAGAAAAATGGATGAATACGTTGCCAGCTTGCACTCT	1744
Db	1636	TAAAGCAGGATGTCTACCTGTGCCCCAAGAAGATGGATGAGTATGTGCCACGCTACACCT	1695
OY	1745	GCCATCATTTGATGCCCCACCTTACAGAGCTGACAGATGACCAACAAAATATCTGGGACT	1804
Db	1696	GCCTACCTTTGATGCCCACCTTGACAGAGCTGACAGATGAACAGGCCAAAGTATCTGGGACT	1755
OY	1805	CAACAAAAATGGCCATTCAACCTAATTATTACAGATACATAATGGACCATACTACCAAG	1864
Db	1756	CAACAGAAGATGGGCCCTTCAAGCCTAATTACTACAGGATTAAGTTCCTGTAACTCAAAC	1815
OY	1865	GACCACT 1871	
Db	1816	CAGAAATT 1822	

RESULT	15
AR068569	
LOCUS	AR068569 2226 bp DNA linear PAT 29-SEP-1999
DEFINITION	Sequence 2 from patent US 5854023.
ACCESSION	AR068569
VERSION	AR068569.1 GI:6000776
KEYWORDS	
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 2226)
TITLE	Hillman,J.L., Corley,N.C., Lal,P. and Shah,P. Polynucleotides encoding human S-adenosyl-5-homocysteine hydrolase derived from bladder Patent: US 5854023-A 2 29-DEC-1998;
JOURNAL	
FEATURES	Location/Qualifiers
source	1..2226
BASE COUNT	649 a /organism="unknown" 455 c 435 g 675 t 12 others
ORIGIN	

Query Match	31.5%;	Score 806.2;	DB 6;	Length 2226;
Best Local Similarity	99.3%;	Pred. No. 4.7e-153;		
Matches 830;	Conservative	0;	Mismatches 4;	Indels 2;
				Gaps 2;

QY	1728	GTGGCCAGCTTGCATCTGCCATCATTTGATGCCCAACCTTACAGAGCTGACAGATGACCAA	1787
DB	1	GTGGCCAGCTTGCATCTGCCATCATTTGATGCCCAACCTTACAGAGCTGACAGATGACCAA	60
QY	1788	GCAAAATATCTGGGACTCAACAAAATGGGCCATTCAAACTTAATTATACAGATACTAA	1847
DB	61	GCAAAATATCTGGGACTCAACAAAATGGGCCATTCAAACTTAATTATACAGATACTAA	120
QY	1848	TGGACCATTAACCAAGACCAAGTCCACCTGAACCAACACACTCTTAAGAATAATTTTAA	1907

Db	121	TGGACCATACTACCAAGGACCAGCTCCACCCTGAACCACACACTCTTAAGAAAATTTT	180
QY	1908	AGATAACTTTTAATTTTCTTCTTACTCCTTTCCCTTGATTTTTTTCATAATTTCATTC	1967
Db	181	AGATAACTTTTAATTTTCTTCTTACTCCTTTCCCTTGATTTTTTTCATAATTTCATTC	240
QY	1968	TTGTTTTTTCATCTCATTTATCCAAAGTCTCTGCAGACCACACAGAACTGCTCATGGCTC	2027
Db	241	TTGTTTTTTCATCTCATTTATCCAAAGTCTCTGCAGACCACACAGAACTGCTCATGGCTC	300
QY	2028	TTTAGATGAATAGAAAGTTCAGGGTCCCTCACTCTAGTCACTAAAGAAGATTTTACTCC	2087
Db	301	TTTAGATGAATAGAAAGTTCAGGGTCCCTCACTCTAGTCACTAAAGAAGATTTTACTCT	360
QY	2088	CCCAGCCAGAAAGGTGATCTCTCTTTACCATTTCTGGGACTTTAGTCTTAATTAGG	2147
Db	361	CCCAGCCAGAAAGGTGATCTCTCTTTACCATTTCTGGGACTTTAGTCTTAATTAGG	419
QY	2148	TACCTTATTAACAGSAAATGCTAAGGTACCTTCTCTGTGGAACAATCTGCAATGCTTAA	2207
Db	420	TACCTTATTAACAGSAAATGCTAAGGTACCTTCTCTGTGGAACAATCTGCAATGCTTAA	479
QY	2208	TCGCCCTTAAAGAGCCCATTTCTTAGCTGCTGAATCAGTGCTCTTTCACITCTTCAGAG	2267
Db	480	TCGCCCTTAAAGAGCCCATTTCTTAGCTGCTGAATCAGTGCTCTTTCACITCTTCAGAG	539
QY	2268	AAGCAGGAGTGTACCTACCCTGCGCAGGTAGGTTAGATGTGGTGTGCAATGTAATTCC	2327
Db	540	AAGCAGGAGTGTACCTACCCTGCGCAGGTAGGTTAGATGTGGTGTGCAATGTAATTCC	599
QY	2328	CTTAGAAGTTCCAAGCCCTGTTTCTCTGCGTAAGGTGGTATGTCCAGTTCAGAGATGTCT	2387
Db	600	CTTAGAAGTTCCAAGCCCTGTTTCTCTGCGTAAGGTGGTATGTCCAGTTCAGAGATGTCT	659
QY	2388	ATAATGAGCATGGCTGTGTTAAGATCAGAGGCCCACTTGATTTATAGATAGCCCTTCC	2447
Db	660	ATAATGAGCATGGCTGTGTTAAGATCAGAGGCCCACTTGATTTATAGATAGCCCTTCC	719
QY	2448	TCCACTCCACACAGACTGCTCATTTTTCGAGTTTAACTAGACTACACTATTTGAG	2507
Db	720	TCCACTCCACACAGACTGCTCATTTTTCGAGTTTAACTAGACTACACTATA-TTGAG	778
QY	2508	TTTAAATTTTGTCTCTAGGATTTATTTCTGTGTGCCAAAAAATTTTAAAAA	2563
Db	779	TTTAAATTTTGTCTCTAGGATTTATTTCTGTGTGCCAAAAAATTTTAAAAA	834

Search completed: April 20, 2003, 16:28:38
Job time : 6549 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2003, 14:20:39 ; Search time 122 Seconds
(without alignments)
6442.726 Million cell updates/sec

Title: US-09-782-051-1
Perfect score: 2563
Sequence: 1 ggcgcggcgagctgcgagct.....aaaaaaaaaaaaaaaaaaaaa 2563

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/lna/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/lna/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/lna/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/lna/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/lna/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/lna/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	806.2	31.5	2226	2 US-08-896-005-2	Sequence 2, Appl1
2	806.2	31.5	2226	4 US-09-347-878-3	Sequence 3, Appl1
3	406.4	15.9	636	4 US-09-328-111-850	Sequence 850, App
4	314.4	12.3	2211	4 US-09-318-448-26	Sequence 26, Appl
5	314.4	12.3	2211	4 US-09-347-878-2	Sequence 2, Appl1
6	261.4	10.2	292	4 US-09-404-879A-40	Sequence 40, Appl
7	237.8	9.3	441529	2 US-09-103-840A-1	Sequence 1, Appl1
8	227.4	8.9	1812	2 US-08-669-536-1	Sequence 1, Appl1
9	216.6	8.5	1767	3 US-08-930-894-1	Sequence 8, Appl1
10	72.6	2.8	289	1 US-08-204-740-8	Sequence 8, Appl1
11	72.6	2.8	289	3 US-09-081-167A-8	Sequence 8, Appl1
12	72.6	2.8	289	3 US-09-081-395-8	Sequence 8, Appl1
13	72.6	2.8	289	4 US-09-416-833-8	Sequence 8, Appl1
14	72.6	2.8	289	5 PCT-US95-02521-8	Sequence 8, Appl1
15	68.4	2.7	285	1 US-08-204-740-6	Sequence 6, Appl1
16	68.4	2.7	285	3 US-09-081-167A-6	Sequence 6, Appl1
17	68.4	2.7	285	3 US-09-081-395-6	Sequence 6, Appl1
18	68.4	2.7	285	4 US-09-416-833-6	Sequence 6, Appl1
19	68.4	2.7	285	5 PCT-US95-02521-6	Sequence 6, Appl1
20	68.2	2.7	584	4 US-08-998-416-249	Sequence 249, App
21	57.4	2.2	7218	1 US-08-232-463-14	Sequence 14, Appl
22	52.2	2.0	320	4 US-09-165-264-13	Sequence 13, Appl
23	52.2	2.0	320	4 US-09-165-264-7	Sequence 7, Appl1
24	51.6	2.0	319	4 US-09-165-264-8	Sequence 8, Appl1
25	51.6	2.0	7218	1 US-08-232-463-14	Sequence 14, Appl
26	51.2	2.0	320	4 US-09-165-264-14	Sequence 14, Appl
27	50.4	2.0	320	4 US-09-165-264-11	Sequence 11, Appl

C	28	50.2	2.0	152331	3 US-09-128-155-16	Sequence 16, Appl
	29	48.8	1.9	1926	4 US-09-249-585A-2	Sequence 2, Appl1
	30	48.8	1.9	2580	3 US-09-050-863-2	Sequence 2, Appl1
C	31	48.8	1.9	2580	4 US-09-359-081-2	Sequence 1, Appl1
	32	48.8	1.9	5452	2 US-09-130-114-1	Sequence 1, Appl1
	33	48.8	1.9	9600	4 US-08-910-647-1	Sequence 1, Appl1
	34	48.8	1.9	9600	4 US-09-620-925-1	Sequence 1, Appl1
	35	48.8	1.9	10596	1 US-07-884-811-15	Sequence 15, Appl
	36	48.8	1.9	10596	1 US-07-885-971-15	Sequence 15, Appl
	37	48.8	1.9	10596	1 US-08-087-783A-15	Sequence 15, Appl
	38	48.8	1.9	10596	1 US-08-194-088B-15	Sequence 15, Appl
	39	48.8	1.9	10596	2 US-08-194-087-15	Sequence 15, Appl
	40	48.8	1.9	10596	5 PCT-US93-04648-15	Sequence 15, Appl
	41	48.4	1.9	318	4 US-09-165-264-12	Sequence 12, Appl
	42	47.6	1.9	1969	4 US-09-098-628-1	Sequence 1, Appl1
C	43	47.4	1.8	1926	4 US-09-249-585A-4	Sequence 4, Appl1
C	44	47.4	1.8	1931	2 US-09-130-114-2	Sequence 2, Appl1
C	45	46.4	1.8	12001	1 US-08-458-568A-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-08-896-005-2
Sequence 2, Application US/08896005
Patent No. 5854023
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Lal, Preeti
APPLICANT: Shah, Purvi
TITLE OF INVENTION: S-ADENOSYL-L-HOMOCYSTEINE HYDROLASE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/896,005
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0337 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2226 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BLADTUT04
CLONE: 1519044
US-08-896-005-2
Query Match 31.5%; Score 806.2; DB 2; Length 2226;
Best Local Similarity 99.3%; Pred. No. 8e-200;
Matches 830; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

```
QY 1728 GTTCCAGCTTGCACTGCGCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAC 1787
      |||||||
DB 1 GTTCCAGCTTGCACTGCGCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAC 60
QY 1788 GCAAAATATCTGGGACTCAACAAAAATGGGCCATCAACCTAATTATTACAGATACTAA 1847
      |||||||
DB 61 GCAAAATATCTGGGACTCAACAAAAATGGGCCATCAACCTAATTATTACAGATACTAA 120
QY 1848 TGGACCATACTACCAAGAGACAGCCAGTCCAGCTGAACACACACTCTAAGAATATTTTAA 1907
      |||||||
DB 121 TGGACCATACTACCAAGAGACAGCCAGTCCAGCTGAACACACACTCTAAGAATATTTTAA 180
QY 1908 AGATACTTTTATTTCTTCTTACTCCTTCTCTGATTTTCTCTATATTTCAATTC 1967
      |||||||
DB 181 AGATACTTTTATTTCTTCTTACTCCTTCTCTGATTTTCTCTATATTTCAATTC 240
QY 1968 TTGTTTTTCACTCATTTATCCAAAGTCTGACAGACACACAGAACTGCTCATGGCTC 2027
      |||||||
DB 241 TTGTTTTTCACTCATTTATCCAAAGTCTGACAGACACACAGAACTGCTCATGGCTC 300
QY 2028 TTTAGATGAATAGAAAGTTCAGGGTCCCTCCTACTCTAGTCACTAAAGAAGATTTTACTCC 2087
      |||||||
DB 301 TTTAGATGAATAGAAAGTTCAGGGTCCCTCCTACTCTAGTCACTAAAGAAGATTTTACTCT 360
QY 2088 CCCAGCCCCAGAAAGTGATTTCTTCTTACCATTCTGGGACTTTAGTCTTAATTAGG 2147
      |||||||
DB 361 CCCAGCCCCAGAAAGTGATTTCTTCTTACCATTCTGGGACTTTAGTCTTAATTAGG 419
QY 2148 TACCTTATTAACAGGAAATGCTAAGTACCTTCTCTGTGAACAATCTGCAATGCTTAA 2207
      |||||||
DB 420 TACCTTATTAACAGGAAATGCTAAGTACCTTCTCTGTGAACAATCTGCAATGCTTAA 479
QY 2208 TCGCCTTAAAGAGCCCATTTCTTAGCTGCTGAATCAAGTCTTTCACCTCTTCAGAG 2267
      |||||||
DB 480 TCGCCTTAAAGAGCCCATTTCTTAGCTGCTGAATCAAGTCTTTCACCTCTTCAGAG 539
QY 2268 AAGCAGGATGCTACCTACCCGCGAGTAGGTAGATGTTGGTGTGCATGTTAATTTC 2327
      |||||||
DB 540 AAGCAGGATGCTACCTACCCGCGAGTAGGTAGATGTTGGTGTGCATGTTAATTTC 599
QY 2328 CTTAGAGTTCCAAAGCCCTGTTCTCTGCTGAAGGTGATGTCCAGTTCAGAGATGTGT 2387
      |||||||
DB 600 CTTAGAGTTCCAAAGCCCTGTTCTCTGCTGAAGGTGATGTCCAGTTCAGAGATGTGT 659
QY 2388 ATAATGAGCATGGCTGTGTAAGATCAGAGGCCCCACTTGATTTATAGTATAGCCCTTCC 2447
      |||||||
DB 660 ATAATGAGCATGGCTGTGTAAGATCAGAGGCCCCACTTGATTTATAGTATAGCCCTTCC 719
QY 2448 TCCACTCCCAACAGACTTGTCTATTTTTCGAGTTTAACTAGACTACACTCTA-TTGAG 2507
      |||||||
DB 720 TCCACTCCCAACAGACTTGTCTATTTTTCGAGTTTAACTAGACTACACTCTA-TTGAG 778
QY 2508 TTTAATTTTGTCTCTAGGATTTATTTCTGTGCCAAAAAATTTTAAAAA 2563
      |||||||
DB 779 TTTAATTTTGTCTCTAGGATTTATTTCTGTGCCAAAAAATTTTAAAAA 834

RESULT 2
US-09-347-878-3
; Sequence 3, Application us/09347878C
; Patent No. 6376210
; GENERAL INFORMATION:
; APPLICANT: Yuan, Chong
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES
; FILE REFERENCE: 25885-1651
; CURRENT APPLICATION NUMBER: US/09/347,878C
; CURRENT FILING DATE: 1999-07-06
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2226
; TYPE: DNA
```

```
ORGANISM: Homo sapiens
FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2226)
OTHER INFORMATION: Polynucleotide encoding human
OTHER INFORMATION: S-adenosyl-5-homocysteine hydrolase (SAHH) derived
OTHER INFORMATION: from bladder; n-a, c, g, or t
PUBLICATION INFORMATION:
PATENT DOCUMENT NUMBER: 08/896,005
PATENT FILING DATE: 1997-07-17
PUBLICATION DATE: 1998-12-29
US-09-347-878-3
```

```
Query Match 31.5%; Score 806.2; DB 4; Length 2226;
Best Local Similarity 99.3%; Pred. No. 8e-200;
Matches 830; Conservative 0; Mismatches 4; Indels 2; Gaps 2;
```

```
QY 1728 GTTCCAGCTTGCACTGCGCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAC 1787
      |||||||
DB 1 GTTCCAGCTTGCACTGCGCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAC 60
QY 1788 GCAAAATATCTGGGACTCAACAAAAATGGGCCATCAACCTAATTATTACAGATACTAA 1847
      |||||||
DB 61 GCAAAATATCTGGGACTCAACAAAAATGGGCCATCAACCTAATTATTACAGATACTAA 120
QY 1848 TGGACCATACTACCAAGAGACAGCCAGTCCAGCTGAACACACACTCTAAGAATATTTTAA 1907
      |||||||
DB 121 TGGACCATACTACCAAGAGACAGCCAGTCCAGCTGAACACACACTCTAAGAATATTTTAA 180
QY 1908 AGATACTTTTATTTCTTCTTACTCCTTCTCTGATTTTCTCTATATTTCAATTC 1967
      |||||||
DB 181 AGATACTTTTATTTCTTCTTACTCCTTCTCTGATTTTCTCTATATTTCAATTC 240
QY 1968 TTGTTTTTCACTCATTTATCCAAAGTCTGACAGACACACAGAACTGCTCATGGCTC 2027
      |||||||
DB 241 TTGTTTTTCACTCATTTATCCAAAGTCTGACAGACACACAGAACTGCTCATGGCTC 300
QY 2028 TTTAGATGAATAGAAAGTTCAGGGTCCCTCCTACTCTAGTCACTAAAGAAGATTTTACTCC 2087
      |||||||
DB 301 TTTAGATGAATAGAAAGTTCAGGGTCCCTCCTACTCTAGTCACTAAAGAAGATTTTACTCT 360
QY 2088 CCCAGCCCCAGAAAGTGATTTCTTCTTACCATTCTGGGACTTTAGTCTTAATTAGG 2147
      |||||||
DB 361 CCCAGCCCCAGAAAGTGATTTCTTCTTACCATTCTGGGACTTTAGTCTTAATTAGG 419
QY 2148 TACCTTATTAACAGGAAATGCTAAGTACCTTCTCTGTGAACAATCTGCAATGCTTAA 2207
      |||||||
DB 420 TACCTTATTAACAGGAAATGCTAAGTACCTTCTCTGTGAACAATCTGCAATGCTTAA 479
QY 2208 TCGCCTTAAAGAGCCCATTTCTTAGCTGCTGAATCAAGTCTTTCACCTCTTCAGAG 2267
      |||||||
DB 480 TCGCCTTAAAGAGCCCATTTCTTAGCTGCTGAATCAAGTCTTTCACCTCTTCAGAG 539
QY 2268 AAGCAGGATGCTACCTACCCGCGAGTAGGTAGATGTTGGTGTGCATGTTAATTTC 2327
      |||||||
DB 540 AAGCAGGATGCTACCTACCCGCGAGTAGGTAGATGTTGGTGTGCATGTTAATTTC 599
QY 2328 CTTAGAGTTCCAAAGCCCTGTTCTCTGCTGAAGGTGATGTCCAGTTACAGATGTGT 2387
      |||||||
DB 600 CTTAGAGTTCCAAAGCCCTGTTCTCTGCTGAAGGTGATGTCCAGTTACAGATGTGT 659
QY 2388 ATAATGAGCATGGCTGTGTAAGATCAGAGGCCCCACTTGATTTATAGTATAGCCCTTCC 2447
      |||||||
DB 660 ATAATGAGCATGGCTGTGTAAGATCAGAGGCCCCACTTGATTTATAGTATAGCCCTTCC 719
QY 2448 TCCACTCCCAACAGACTTGTCTATTTTTCGAGTTTAACTAGACTACACTCTAATTGAG 2507
      |||||||
DB 720 TCCACTCCCAACAGACTTGTCTATTTTTCGAGTTTAACTAGACTACACTCTA-TTGAG 778
QY 2508 TTTAATTTTGTCTCTAGGATTTATTTCTGTGCCAAAAAATTTTAAAAA 2563
      |||||||
DB 779 TTTAATTTTGTCTCTAGGATTTATTTCTGTGCCAAAAAATTTTAAAAA 834
```

```

RESULT 3
US-09-328-111-850/c
; Sequence 850, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steilmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dertl, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 850
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(636)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-111-850

Query Match      15.9%; Score 406.4; DB 4; Length 636;
Best Local Similarity 99.5%; Pred. No. 3.7e-96;
Matches 418; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 707 TACACACATCACAGCCCCAGACAGC-GGTGTGATTGAGACACTCTGCCCCCTGGGGGCTC 765
      |||||||
Db 420 TACACACATCACAGCCCCAGACAGCGGGGTGTGATTGAGACACTCTGCTCGGGGCTC 361

QY 766 AGTGGCCGTGCTGCTGTGAACATCTACTCACTCAGAAATGAAGTAGCTGCAGCACTGG 825
      |||||||
Db 360 AGTGGCCGTGCTGCTGTGAACATCTACTCACTCAGAAATGAAGTAGCTGCAGCACTGG 301

QY 826 CTGAGGCTGAGTTGCAGTGTTCGCTTGGAAGGGCGAGTCAGAAAGATGACTTCTGTGTGT 885
      |||||||
Db 300 CTGAGGCTGAGTTGCAGTGTTCGCTTGGAAGGGCGAGTCAGAAAGATGACTTCTGTGTGT 241

QY 886 GATTTGACCGCTGTGTGAACATGATGGGTGGCAGGCCAACATGATCTCTGATGATGGGG 945
      |||||||
Db 240 GATTTGACCGCTGTGTGAACATGATGGGTGGCAGGCCAACATGATCTCTGATGATGGGG 181

QY 946 GAGACTTAACCCACTGGGTTTAAAGAGATATCCAAACGTGTTTAAAGAGATCCGAGGCA 1005
      |||||||
Db 180 GAGACTTAACCCACTGGGTTTAAAGAGATATCCAAACGTGTTTAAAGAGATCCGAGGCA 121

QY 1006 TTTGGAAGAGAGCGTGAAGTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGC 1065
      |||||||
Db 120 TTTGGAAGAGAGCGTGAAGTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGC 61

QY 1066 TCTGTGTTCCGGCCATGAACGTCAATGATTTCTGTATACCAACAGAGATTGATACTGT 1125
      |||||||
Db 60 TCTGTGTTCCGGCCATGAACGTCAATGATTTCTGTATACCAACAGAGATTGATACTGT 1

```

```

; Patent No. 6210950
; GENERAL INFORMATION:
; APPLICANT: Johnson, William G.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; FILE REFERENCE: 601-1-057
; CURRENT APPLICATION NUMBER: US/09/318,448
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; LENGTH: 2211
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-318-448-26

Query Match      12.3%; Score 314.4; DB 4; Length 2211;
Best Local Similarity 61.9%; Pred. No. 6.1e-72;
Matches 498; Conservative 0; Mismatches 306; Indels 0; Gaps 0;

QY 580 ACATCAAGCAGGCGAGAATTTGACGCCGGGAGATTGAGATTGCAGACCAAGACATGTCTG 639
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 79 ACATCGCCCTGGCTGCCCTGGGGAGCCGACGCCCTTGACATTTGCTGAGAACGAGATGCCG 138
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 640 CTCTGATTTCACTCAGGAACGTCGTCAGGGGGAGAGCCCTTGCTGCTCTAAATAG 699
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 139 GCCGTGATGCGTATGCGGAGCGGTACTCGGCCTCCAGCCACTGAAGGGCCCGCATCG 198
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 700 TGGGCTGTACACACATCACAGCCACAGACGCGGTGTTGATTGAGACACTCTGTGCCCTGG 759
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 199 CTGGCTGCCTGCACATGACCGGTGAGAGCGCCGCTCCTCATTGAGACCCCTGTCACCCCTGG 258
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 760 GGGCTCAGTGCCCGCTGGTCTGCTGTGTAACATCTACTCACTCAGAATGAAGTAGCTGCAG 819
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 259 GTGTGAGGTGCAAGTGTGTCAGCTGCACATCTTCTCCACCAGAACCATGCGCGGGCTG 318
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 820 CACTGCTGAGGCTGGAGTTGCACTGTCCTTGGAAGGGCGAGTCAGAGATGACTTCT 879
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 319 CCATTGCCAAGGCTGGCATTCGCGTGTATGCTTGAAGGGCGAAGCGAGAGATACC 378
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 880 GGTGTGTATGTACCGCTGTGTGAACATGATGGGTGGCAGGCCCAACATGATCCTGGATG 939
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 379 TGTGTGCATGTAGCAGAGACCCCTGTACTCAAGAGAGCGGCCCTCAACATGATTTGAGAG 438
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 940 ATGGGGAGACTTAAACCACCTGGGTATTAAGAAGTATCCAAACGTGTTTAAGAAGATCC 999
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 439 ACCGGGGCGACCTCACCAACCTCATCCACCAAGTATCCCGCAGCTTCTGCCAGGCATCC 498
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1000 GAGGCATGTGGAAGAGAGCGGTGACTGGTGTTCACAGGCTGTATCAGTCTCCAAAGCTG 1059
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 499 GAGGCATCTGTAGAGAGACACGACTGGGGTCCACAACCTTACCAAGATGATGGCCATG 558
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1060 GGAAGCTGTGTGTTCCGGCATGAACGTCAATGATTTCTGTTACCAAAAGAGTTGATA 1119
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 559 GGATCCTCAAGGTGCTGCTGCATCAATGATGAATGACTCCGTCCACCAAGAGCAAGTTGACA 618
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1120 ACTGTACTGCTGCCGAGAAATCCATTTTGGATGGCCCTGAAGAGACACAGATGTGATGT 1179
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 619 ACCCTATGCGCTGCCGGGAGTCCCTCATAGATGGCATCAAGCGGGCCACAGATGTGATGA 678
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1180 TTGTGGAACAACAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1239
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 679 TTGCGCGAAGTAGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 738
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1240 CTCTCAAGCTCTTGAGCAATTTGTACTTACCGAAATGACCCCATCTGTCTGTGC 1299
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 739 CCCTGCGGGGTTTCGAGCCCGCGCTCATCATCACCGAGATTGACCCCATCAACGCACTGC 798
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1300 AGGCTGCATGAGTGGGTTCAGGGTGTGTAAAGCTAATGAAGTCATCCGGCAAGTGCATG 1359
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 799 AGGCTGCATGAGGGGCTATGAGGTGACCAACCATGATGAGGCCCTGTCAAGAGGGCAAGA 858

```

QY 1360 TCGTAATACTTGACAGGAATA 1383
DB 859 TCTTTGTACACACAGGCTGTA 882

RESULT 5

US-09-347-878-2
; Sequence 2, Application US/09347878C
; Patent No. 6376210
; GENERAL INFORMATION:
; APPLICANT: Yuan, Chong
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES
; FILE REFERENCE: 25885-1651
; CURRENT APPLICATION NUMBER: US/09/347, 878C
; CURRENT FILING DATE: 1999-07-06
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2211
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Human S-adenosylhomocysteine hydrolase cDNA
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: M61831/GenBank
US-09-347-878-2

Query Match 12.3%; Score 314.4; DB 4; Length 2211;
Best Local Similarity 61.9%; Pred. No. 6.1e-72;
Matches 498; Conservative 0; Mismatches 306; Indels 0; Gaps 0;

QY 580 ACATCAAGCAGGAGAAATTTGGAGCGCGGAGATGAGATTGACAGCAAGACATGCTG 639
DB 79 ACATCGGCTGCTGCTGCTGGGAGCGCAAGCCCTGACATGCTGAGAACGAGATGCCGG 138
QY 640 CTCTGATTTCACTCAGGAACGCTCTCAGGGGAGAGAGCCCTTGCTGCTAAATAG 699
DB 139 GCGTATGCTGATGCGGAGCGGTAATCTGCGCTCCAGCCACTGAAGGGCGCCGATCG 198
QY 700 TGGGCTGTACACACATCACAGCCAGACAGCGGTGTTGATTGAGACACTCTGCTGG 759
DB 199 CTGGCTGCTGACATGACCGGTGAGAGCGCCCTCTCATGTGAGACCTCGTACCCCTG 258
QY 760 GGGCTCAGTCCGCTGCTGCTGTAACATCTACTCAACTCAGATGAAGTAGCTGAC 819
DB 259 GTGCTGAGGTGAGTGTCCAGCTGCAACATCTCTCCACCCAGAACCATGCGCGCTG 318
QY 820 CACTGCTGAGGCTGAGTGTGAGTGTGCTTGAAGGGCGAGTCAAGAGATGACTCT 879
DB 319 CCATTGCCAAGGCTGGCATTCGGTGTATGCTGGAAGGGCGAAGCGAGGAGTACC 378
QY 880 GGTGCTGATTTGACCGCTGTTGACATGATGGGTGGCAGGCCAATGATCTGATG 939
DB 379 TGTGCTGATTTGACAGACCTGTAATCAAGGAGCGCCCTCAACATGATTTGAGC 438
QY 940 ATGGGGAGACTTAACCCACTGGTTTATAAGAGATATCCAAACGTTTAAAGATCC 999
DB 439 ACGGGGGGAGCTCAACCACTCAACCAAGTACCCGAGCTTCTGCCAGGATCC 498
QY 1000 GAGGCAATTTGGAAGAGAGCGTGTGCTCAGAGGCTGTATCAGCTCTCCAAAGCTG 1059
DB 499 GAGGCAATTTGGAAGAGAGCGTGTGCTCAGAGGCTGTATCAGCTCTCCAAAGCTG 1059
QY 1060 GGAAGCTGTGTGCTGCGGCTGTAACGTAATGATTTCTTACCAAGAGATTGATA 1119
DB 559 GGAATCTCAAGGTGCTGCTGCTCAATGTAATGACTCCGTACCAAGAGCAAGTTGACA 618
QY 1120 ACTGTACTGCTGCGGAGATCAATTTGATGCGCTGAAGAGACACAGATGTGATG 1179
DB 619 ACCCTATGCTGCGGAGAGCTCCATAGATGATGATCAAGCGGCGACAGATGTGATA 678
QY 1180 TTGGTGGAAACAAGTGTGTGTGCTATGATGATGAGTAGGAGGCTGCTGTGCTG 1239

DB 679 TTCCGGCAGAGTAGCGGTGTAGCAGGCTATGATGTGGCAAGGGCTGTGCCAGG 738
QY 1240 CTCTCAAGCTCTTGAGCAATTTCTACATTTACCGAAATCGACCCCATCTGTCTG 1299
DB 739 CCCTGCGGGTTCGAGCGCCGCTCATCATCACCAGATGACCCCATCAACGACTGC 798
QY 1300 AGGCTGATGATGGGTTCAGGGTGTAAAGCTAAATGAATGATCGGGCAAGTCATG 1359
DB 799 AGGCTGATGATGGGTTCAGGGTGTAAAGCTAAATGAATGATCGGGCAAGTCATG 858
QY 1360 TCGTAATACTTGACAGGAATA 1383
DB 859 TCTTTGTACACACAGGCTGTA 882

RESULT 6

US-09-404-879A-40/c
; Sequence 40, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404, 879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 40
; LENGTH: 292
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-404-879A-40

Query Match 10.2%; Score 261.4; DB 4; Length 292;
Best Local Similarity 99.3%; Pred. No. 1.2e-58;
Matches 273; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 2261 TTCAGAGAGCAGGATGTTACCTACCCGCGCAGGTAGGTAGATGGGTGTCATGT 2320
DB 282 TTCAGGAGCAGGATGTTACCTACCCGCGCAGGTAGGTAGATGGGTGTCATGT 223
QY 2321 AATTCCCTTAGAAGTTCAGGCTGTTCTGCGTAAAGGTGATGTCAGTTCAGA 2380
DB 222 AATTCCCTTAGAAGTTCAGGCTGTTCTGCGTAAAGGTGATGTCAGTTCAGA 163
QY 2381 GATGTGATATGAGCATGGCTGTTAAGATCAGAGGCCACTGATTTATAGTATAG 2440
DB 162 GATGTGATATGAGCATGGCTGTTAAGATCAGAGGCCACTGATTTATAGTATAG 103
QY 2441 CCCTTCTCCACTCCACACAGCTTCTCATTTTTCGAGTTTAACTAGACTACACTCT 2500
DB 102 CCCTTCTCCACTCCACACAGCTTCTCATTTTTCGAGTTTAACTAGACTACACTCT 43
QY 2501 ATTTGAGTTAATTTTGTCTCTTAGATTTATTTTC 2535
DB 42 A-TTGAATTAAATTTTGTCTCTTAGATTTATTTTC 9

RESULT 7

US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00

```

; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37RV
US-09-103-840A-1

```

Query Match	9.38;	Score 237.8;	DB 4;	Length 4411529;
Best Local Similarity	56.68;	Pred. No. 3e-50;		
Matches 503; Conservative	0;	Mismatches 377;	Indels 9;	Gaps 3;

QY 970 AGAAGTATCCAAACGCTGTTTAAAGAAGATCCGAGGCACTGTGTGGAAGAAGAGCGGTGACTGGTG 1029
 |||||
 Db 3629036 ACAAGTGGACCACAGATAGCCGAGTCGGTCAAGGGCGCTCACCGAGGAGAACCCACCGGCG 3628977
 QY 1030 TTCACAGGCTGTATCACTCTCCAAAGCTGGGAAGCTCTGTGTCTCCGGCATGAACGTCA 1089
 |||||
 Db 3628976 TGTGCGGCTCTACCAATTCGCCCGCGGGGATCTGGGCTTCCCGGCGATCAACGTCA 3628917
 QY 1090 ATGATTTCTGTACCAACAAGAAGTTTGATACTTGTACTGTCTGCTCCGAGAATCCATTTTGG 1149
 |||||
 Db 3628916 ACGACTCGGTGACCAAGTCCAAATTCGACACACAAGTACGGCACTCCGCACTCCCTGATCG 3628857
 QY 1150 ATGGCCCTGAAGAGGACCACAGATGTGATGTTTGGTGGGAACAAGTGGTGTGTGGCT 1209
 |||||
 Db 3628856 ACGGCATCAACCGCGGACCGGACCGGCTGATCGCGGTAAGAAGTCTCATCTGCGGCT 3628797
 QY 1210 ATGCTGAGGTAGGCAAGGCGCTGTGCTGTCTCTCAAGCTCTTGAGCAATTGTCTACA 1269
 |||||
 Db 3628796 ACGCGACGTCGGTAAGGCTGTGCGGAGCGATGAAGGGCCAGGAGCGCGGTCTCCG 3628737
 QY 1270 TTACCGAAATCGACCCCATCTGTGCTCTGCAAGCCCTGCATGGAATGGGTTCAAGGTGTAA 1329
 |||||
 Db 3628736 TCACCGAGATCGACCCGATCAACGCGCTGCAAGCCATGATGGAGGCTTCGACGTGTC 3628677
 QY 1330 AGCTAATGAAGTCATCCGCGAAGTCGATGTCGTAATTAATTGACACAGGAATAAGAATG 1389
 |||||
 Db 3628676 CCGTCGAGGAGGCCATCGGSGACGCCGACATCGTCTAACCGCGAACCGGCAACAAGACA 3628617
 QY 1390 TAGTGACACGGGAGCACTTGGATCGCATGAAAAACAGTTGTATCGTATGCAATATGGGCC 1449
 |||||
 Db 3628616 TCATCATGCTCGAGCACATTAAGCGATGAAGAGCACACGCGATCTGGAAATATCGGCC 3628557
 QY 1450 ACTCCAACACAGAAATGATGTGACCAAGCCTCCGCACTCCGAGCTGACGTGGAGCGAG 1509
 |||||
 Db 3628556 ACTTCGACACAGAGATCGACATGGCCGGGCTGAGCGCTCCGGGGGACACAGGGTCAACG 3628497
 QY 1510 TACGTTCTCAGGTGAGCAATGTCACTCTGGCCAGA--TGGCAACGAGTTGTCTCTCTGG 1566
 |||||
 Db 3628496 TCAAGCCTCAGGTGACCTGTGGACCTTTGGCGACACGGGCGGCTCGATCATCTGTCTGT 3628437
 QY 1567 CAGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTTCCACC--TTGTCTCTGTTCCA 1623
 |||||
 Db 3628436 CCGAGGGGCGGCTGTGAACTGGGCAATGCCACCGGGCACCCCTCTGTCTGATGAGCA 3628377
 QY 1624 TCACAGCCACACACAGGCTTTGGCACTGATAGAACTCTAATATGCAACCCGAGGGCGAT 1683
 |||||
 Db 3628376 AAGGCTTCGCTAACACAGACGATTCGCCAGATCGAGCTGTGGACCAAGAAGACGAGAG--T 3628320
 QY 1684 ACAAGCAGGATGTGTACTTGTCTTCCCTAAGAAATGGATGAATACGTGCGCAGCTGCAATC 1743
 |||||
 Db 3628319 ACGACAAACGAGGTGTACCGGCTGCCCAAGCACCTCGACGAGAAGGTGGCTCGAATCCATG 3628260
 QY 1744 TGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGAC 1803
 |||||
 Db 3628259 TCGAGGCCCTTGGCGGTCACTGACCAAGCTGACCAAGGAGCAGGCCGAATACCTGGGCG 3628200
 QY 1804 TCACAATAATGGGCGCATTCAAACCTAATTATTACAGATACTAATGGAC 1852
 |||||

RESULT 8
US-08-669-536-1

; Sequence 1, Application US/08669536
; Patent No. 5910444

GENERAL INFORMATION:

APPLICANT: MASUTA, CHIKARA

APPLICANT: UEHARA, KYOKO
APPLICANT: TANAKA, HIDEO

APPLICANT: KIMATA SHIGERU
; IANAKA, HIDEO
:

TITLE OF INVENTION: ORGANIC

TITLE OF INVENTION: S-ADENOSYLHOMOCYSTEINE HYDROLASE GENE

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH

STREET: PO BOX 747
CITY: BAITA QUNDUQ

CITY: FALLS CHURCH
STATE: VA

STATE.
 COUNTY.

ZIP: 22040-0

COMPUTER READABLE

MEDIUM TYPE: F1

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS

SOFTWARE: Patentin Release
; CREDITUM "PATENTIN CAME"

CURRENT APPLICATION DATA:
APPLICATION NUMBER: TIS/009/6660 536

APPLICATION NUMBER: 03/06/003,330

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: MURPHY JR, GE

REGISTRATION NUMBER: 28,

REFERENCE/DOCKET NUMBER:

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 305-8000

TELEPHONE: (703) 203-8000
TELEFAX: (703) 205-8050

INFORMATION FOR SEO ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1812 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA CO
HS-08-6669-536-1

C
U
C
C
U
U
U
C
H

Query Match 8.98;

Best Local Similarity 56.18;

Matches 514; Conservative

Query Match	8.9%;	Score 227.4;	DB 2;	Length 1812;
Best Local Similarity	56.1%;	Pred. No. 2.3e-49;		
Matches 514;	Conservative 0;	Mismatches 391;	Indels 12;	Gaps 4;

QY	972	AAGTATCCAAACGTGTTTAA	GAGATCCGAGGCATTGTGGA	GAGAGAGCGCTGACTGCTG	TT	1031
Db	634	AAATATACCAAGATGAAGAA	GAAGACTCGTCGGTGTTC	TGAGGAAC	TACCACTGGA	TT 693
QY	1032	CACAGCGCTGATCAGCTCT	CCAAGCTGGGAAGCTCTGT	TCCGGCCATG	GAACGTCAAT	1091
Db	694	AAGAGGCTTATCAGATGCA	GAGCTAATGGAAC	TTGCTTTCCCTGCTATTA	TGTTAAT	753
QY	1092	GATTCCTGTACCAACAGA	AGTTGATACTGTACTGCTG	CCGAGAACTTTTG	CAAT	1151
Db	754	GATTCGTACCAAGACAGCA	AGTTGCAACA	CTGTACGATGCGC	CAC	TCACTGCCGAT 813
QY	1152	GGCCTGAAGAGCACAGAT	GTGATGTTGGTGGAAACA	AGTGTGTGTGTG	GGCTAT	1211
Db	814	GGTCTCATGAGGGCTACT	GATGTTATGATTGCCG	AAAGTTGCCCTGTG	CTGGTTAT	873
QY	1212	GGTAGGATGAGCAAGGCT	GCTGCTGCTCTCAAA	GCCTCTTGAGAGCA	ATTGCTACAT	TT 1271
Db	874	GGAGATGTCCGCCAAGG	GTGTGCTGCTGCTTGA	AAACAAGCCGCTG	CCCGTGTGATTGTG	933
QY	1272	ACCGAAATCGACCCCAT	CTGTGCTCTGCAAGC	CTGCATGGATGGGT	TCAGGGTGGTAAAG	1331

Db 1316 AAGTATGAGAGAGAGTTACGTGCTCCCAAGCATCTGTGATGAGAAAGTAGAGCGCTT 1375
QY 1740 CATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTG 1799
Db 1376 CACTTGGGCAAGCTCGGAGCCAAAGCTTACAAAGCTCAGCCCTTCACAGGCGGAGTACATC 1435
QY 1800 GGACTCAACAAAATGGCCATTCAACTAATTTACAGATCTA 1846
Db 1436 AGCGTCCCATCGAGGGTCCCTACAGCCACTCAGTACAGGTACTA 1482

RESULT 10

US-08-204-740-8

Sequence 8, Application US/08204740

Patent No. 5753432

GENERAL INFORMATION:

APPLICANT: Gudkov, Andrei

APPLICANT: Kazarov, Alexander

APPLICANT: Mazo, Ilya

APPLICANT: Roninson, Igor B.

TITLE OF INVENTION: Methods for Identifying Genetic

TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant

TITLE OF INVENTION: Growth in Cancer Cells

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: Allgretti & Witcoff, Ltd.

STREET: 10 S. Wacker Drive, Suite 3000

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/204,740

FILING DATE: 04-MAR-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: NO. 5753432nan, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 93,354-C

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-715-1000

TELEFAX: 312-715-1234

TELEX: 910-221-5317

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 289 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-08-204-740-8

Query Match 2.8%; Score 72.6; DB 1; Length 289;
Best Local Similarity 58.6%; Pred. No. 1.5e-09;
Matches 126; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 580 ACATCAAGCAGCAGAAATTGGACGCGGAGATTGAGATTGACAGCAAGACATGCTG 639
Db 75 ACATCGGCGCTGGCTGGCTGGGAGCGAGCGCCCTGGACATTTGCTGAGAAGAGATGCCG 134
QY 640 CTCTGATTTCACTCAGAAACGTGCTCAGGGGGGAGAGCCCTTGCTGCTAAATAG 699
Db 135 GCCTGATGCGTAATGCGGAGCGGTACTCGGCTCCAGCCACTGAAGGGCGCCGATCG 194
QY 700 TGGCTGTACACATCAGACCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGG 759
Db 195 CTGCTGCTGCTGACATGACCGGTGAGAGACGCGCCCTCTCATTTGAGACCCCTGTACCCCTGG 254

QY 760 GGGCTCAGTGGCGCTGCTGCTGTGTAACATCTAC 794
Db 255 GTGCTGAGGTGAGTGGTCCAGCTGCAACATCTTC 289

RESULT 11

US-09-081-167A-8

Sequence 8, Application US/09081167A

Patent No. 6083745

GENERAL INFORMATION:

APPLICANT: Gudkov, Andrei

APPLICANT: Kazarov, Alexander

APPLICANT: Mazo, Ilya

APPLICANT: Roninson, Igor B.

TITLE OF INVENTION: Methods for Identifying Genetic

TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant

TITLE OF INVENTION: Growth in Cancer Cells

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff

STREET: 300 S. Wacker Drive, 32nd Floor

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/081,167A

FILING DATE: 18-MAY-1998

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: NO. 6083745nan, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 93,354-KK

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-913-0001

TELEFAX: 312-913-0002

TELEX:

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 289 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-09-081-167A-8

Query Match 2.8%; Score 72.6; DB 3; Length 289;
Best Local Similarity 58.6%; Pred. No. 1.5e-09;
Matches 126; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 580 ACATCAAGCAGCAGAAATTGGACGCGGAGATTGAGATTGACAGCAAGACATGCTG 639
Db 75 ACATCGGCGCTGGCTGGCTGGGAGCGAGCGCCCTGGACATTTGCTGAGAAGAGATGCCG 134
QY 640 CTCTGATTTCACTCAGAAACGTGCTCAGGGGGGAGAGCCCTTGCTGCTAAATAG 699
Db 135 GCCTGATGCGTAATGCGGAGCGGTACTCGGCTCCAGCCACTGAAGGGCGCCGATCG 194
QY 700 TGGCTGTACACATCAGACCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGG 759
Db 195 CTGCTGCTGCTGACATGACCGGTGAGAGACGCGCCCTCTCATTTGAGACCCCTGTACCCCTGG 254
QY 760 GGGCTCAGTGGCGCTGCTGCTGTGTAACATCTAC 794
Db 255 GTGCTGAGGTGAGTGGTCCAGCTGCAACATCTTC 289

RESULT 12
US-09-081-395-8
; Sequence 8, Application US/09081395
; Patent No. 6083746
; GENERAL INFORMATION:
; APPLICANT: Gudkov, Andrei
; APPLICANT: Kazarov, Alexander
; APPLICANT: Mazo, Ilya
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods for Identifying Genetic
; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
; TITLE OF INVENTION: Growth in Cancer Cells
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 S. Wacker Drive, 32nd Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/081,395
; FILING DATE: 18-MAY-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6083746nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 93,354-KK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; TELEX:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 289 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; US-09-081-395-8
Query Match 2.8%; Score 72.6; DB 3; Length 289;
Best Local Similarity 58.6%; Pred. No. 1.5e-09;
Matches 126; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
QY 580 ACATCAAGCAGCAGAATTGGACGCCGGAGATTGAGATTGCAGAGCAAGACATGCTG 639
DB 75 ACATCGCCCTGGCTGCGGGAGCGGAGCCCTGGACATTGCTGAGAGCGAGATGCCGG 134
QY 640 CTCTGATTTCACTCAGAAACGTGCTCAGGGGGAGAGCCCTTGCTGCTAAATAG 699
DB 135 GCCTGATGCGTATGCGGAGCGGTACTCGGCTCCAGCCACTGAAGGGCGCCGCATCG 194
QY 700 TGGGCTGTACACATCAGACGCCAGACGGGTGTGATTGAGACACTCTGTGCCCTGG 759
DB 195 CTGGCTGCGCTGCACATGACCGGTGAGAGCGCCGCTCTCATTTGAGAGCCCTGTCACCCCTGG 254
QY 760 GGGCTCAGTGGCCGCTGCTGCTGTGTAACATCTAC 794
DB 255 GTGCTGAGGTGAGTGTCCAGCTGCAACATCTTC 289
RESULT 13
US-09-416-833-8
; Sequence 8, Application US/09416833
; Patent No. 6197521
; GENERAL INFORMATION:
; APPLICANT: Gudkov, Andrei

APPLICANT: Kazarov, Alexander
APPLICANT: Mazo, Ilya
APPLICANT: Roninson, Igor B
TITLE OF INVENTION: Methods for Identifying Genetic
TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
TITLE OF INVENTION: Growth in Cancer Cells
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Allgreti & Witcoff, Ltd.
STREET: 10 S. Wacker Drive, Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/416,833
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/204,740
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: No. 6197521nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 93,354-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1234
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 289 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-416-833-8
Query Match 2.8%; Score 72.6; DB 4; Length 289;
Best Local Similarity 58.6%; Pred. No. 1.5e-09;
Matches 126; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
QY 580 ACATCAAGCAGCAGAATTGGACGCCGGAGATTGAGATTGCAGAGCAAGACATGCTG 639
DB 75 ACATCGCCCTGGCTGCGGGAGCGGAGCCCTGGACATTGCTGAGAGCGAGATGCCGG 134
QY 640 CTCTGATTTCACTCAGAAACGTGCTCAGGGGGAGAGCCCTTGCTGCTAAATAG 699
DB 135 GCCTGATGCGTATGCGGAGCGGTACTCGGCTCCAGCCACTGAAGGGCGCCGCATCG 194
QY 700 TGGGCTGTACACATCAGACGCCAGACGGGTGTGATTGAGACACTCTGTGCCCTGG 759
DB 195 CTGGCTGCGCTGCACATGACCGGTGAGAGCGCCGCTCTCATTTGAGAGCCCTGTCACCCCTGG 254
QY 760 GGGCTCAGTGGCCGCTGCTGCTGTGTAACATCTAC 794
DB 255 GTGCTGAGGTGAGTGTCCAGCTGCAACATCTTC 289
RESULT 14
PCT-US95-02521-8
; Sequence 8, Application PC/TUS9502521
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Methods for Identifying Genetic
; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
; TITLE OF INVENTION: Growth in Cancer Cells
; NUMBER OF SEQUENCES: 13

QY	1728	GTGGCCAGCTTGCATCTGCCATCATTTGATGGCCACCTTACAGAGCTGCACAGATGACCAA	1787
Db	1	GTGGCCAGCTTGCATCTGCCATCATTTGATGGCCACCTTACAGAGCTGCACAGATGACCAA	60
QY	1788	GCAAAATATCTGGGACTCAACAAAAATGGGCCATTCAACTTAATTAAGATACTAA	1847
Db	61	GCAAAATATCTGGGACTCAACAAAAATGGGCCATTCAACTTAATTAAGATACTAA	120
QY	1848	TGGACCATACTACCAAGGACCAGTCCACCTGAACACACACACTCTAAGAATATTTT	1907
Db	121	TGGACCATACTACCAAGGACCAGTCCACCTGAACACACACACTCTAAGAATATTTT	180
QY	1908	AGATACTTTTATTTTCTTCTTCTACTCCTTCCCTGTGATTTTTTCTATAATTCTATC	1967
Db	181	AGATACTTTTATTTTCTTCTTCTACTCCTTCCCTGTGATTTTTTCTATAATTCTATC	240
QY	1968	TTGTTTTTTCATCTCATTTATCCAAGTCTGCAGAGCCACACAGAACTTGCTTCATGGCTC	2027
Db	241	TTGTTTTTTCATCTCATTTATCCAAGTCTGCAGAGCCACACAGAACTTGCTTCATGGCTC	300
QY	2028	TTTAGATGAATAGAAAGTTCCAGGGTCCCTCCTACTCTAGTCACTAAGAAGGATTTTCTCC	2087
Db	301	TTTAGATGAATAGAAAGTTCCAGGGTCCCTCCTACTCTAGTCACTAAGAAGGATTTTCTCT	360
QY	2088	CCCAAGCCAGAAAGGTGATTTCTTCTTTACCATTTCTGGGGACTTAGTCTTAATTAGG	2147
Db	361	CCCAAGCCAGAAAGGTGATTTCT - TCTTACCATTTCTGGGGACTTAGTCTTAATTAGG	419
QY	2148	TACCTTATTAACAGGAATGCTAAGGTACCTTCTCTGTGAAACAATCTGCAATGTCTAA	2207
Db	420	TACCTTATTAACAGGAATGCTAAGGTACCTTCTCTGTGAAACAATCTGCAATGTCTAA	479
QY	2208	TCGCCTTAAAAGAGCCCATTTTCTAGCTGCTGAATCAGTGCCTTTCACTTCTTCAGAG	2267
Db	480	TCGCCTTAAAAGAGCCCATTTTCTAGCTGCTGAATCAGTGCCTTTCACTTCTTCAGAG	539
QY	2268	AAGCAGGATGGTACCTACCCCGCAGGTAGGTAGATGTGGGTGGTGCATGTTAATTCC	2327
Db	540	AAGCAGGATGGTACCTACCCCGCAGGTAGGTAGATGTGGGTGGTGCATGTTAATTCC	599
QY	2328	CTTAGAAGTCCAAAGCCCTGTCTTCTGCGTAAAGGTGTATGTCCAGTTCAGAGATGTGT	2387
Db	600	CTTAGAAGTCCAAAGCCCTGTCTTCTGCGTAAAGGTGTATGTCCAGTTCAGAGATGTGT	659
QY	2388	ATAATGAGCATGGCTTGTTAAGATCAGAGAGCCCACTTGATTTAAGTATAGCCCTTCC	2447
Db	660	ATAATGAGCATGGCTTGTTAAGATCAGAGAGCCCACTTGATTTAAGTATAGCCCTTCC	719
QY	2448	TCCACTCCCAACCAAGACTTGCTCATTTTTCGAGTTTTTAACTAGACTACACTCTATT	2503
Db	720	TCCACTCCCAACCAAGACTTGCTCATTTTTCGAGTTTTTAACTAGACTACACTCTATT	775

RESULT 2
US-09-347-878-3
; Sequence 3, Application US/09347878C

```

: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES
:
: FILE REFERENCE: 25885-1651
:
: CURRENT APPLICATION NUMBER: US/09/347, 878C
:
: CURRENT FILING DATE: 1999-07-06
:
: NUMBER OF SEQ ID NOS: 75
:
: SOFTWARE: PatentIn Ver. 2.0
:
: SEQ ID NO 3
:
: LENGTH: 2226
:
: TYPE: DNA
:
: ORGANISM: Homo sapiens
:
: FEATURE:
:
: NAME/KEY: misc_feature
:
: LOCATION: (1)..(2226)
:

```

```

; OTHER INFORMATION: Polynucleotide encoding human
; OTHER INFORMATION: S-adenosyl-5-homocysteine hydrolase (SAHH) derived
; OTHER INFORMATION: from bladder; n=a, c, g, or t
; PUBLICATION INFORMATION:
; PATENT DOCUMENT NUMBER: 08/896,005
; PATENT FILING DATE: 1997-07-17
; PUBLICATION DATE: 1998-12-29
US-09-347-878-3

```

Query Match	21.68;	Score 553;	DB 4;	Length 2226;
Best Local Similarity	99.68;	Pred. No. 1.4e-236;		
Matches 773; Conservative	0;	Mismatches 2;	Indels 1;	Gaps 1;

OY	1728	GTTGCCAGCTTGCATCTGCCCATCATTTGATGCCACCCTTACAGAGCTGCACAGATGACCAA	1787
Db	1	GTGGCCAGCTTGCACTCGCCATCATTTGATGCCACCCTTACAGAGCTGCACAGATGACCAA	60
OY	1788	GCAAAATACTGGGACTCAACA AAAATGGGCCATTCAACCTTAATTTACAGATACTAA	1847
Db	61	GCAAAATACTGGGACTCAACA AAAATGGGCCATTCAACCTTAATTTACAGATACTAA	120
OY	1848	TGGACCACTACTACCAAGGACCAGTCACCTGAACCACACACTCTAAGAATAATTTTTTA	1907
Db	121	TGGACCACTACTACCAAGGACCAGTCACCTGAACCACACACTCTAAGAATAATTTTTTA	180
OY	1908	AGATACTTTTATTTTCTTCTTACTCCTTCTCCTTGATTTTTTCCATAATTTTCATT	1967
Db	181	AGATACTTTTATTTTCTTCTTACTCCTTCTCCTTGATTTTTTCCATAATTTTCATT	240
OY	1968	TTGTTTTTTCATCTCATTTATCCAAGTTCTGCAGACACACAGAACCTTGCTTCATGGCT	2027
Db	241	TTGTTTTTTCATCTCATTTATCCAAGTTCTGCAGACACACAGAACCTTGCTTCATGGCT	300
OY	2028	TTTAGATGAATAGAAAGTTCAGGGTCCCTCACCTTAGTCACCTAAAGAGATTTTACTCC	2087
Db	301	TTTAGATGAATAGAAAGTTCAGGGTCCCTCACCTTAGTCACCTAAAGAGATTTTACTCT	360
OY	2088	CCCAGCCCCAGAAAGGTGATTTCTTCTTACCATTTCGGGACTTATAGTCTTAATTAGG	2147
Db	361	CCCAGCCCCAGAAAGGTGATTTCT - TCTTACCATTTCGGGACTTATAGTCTTAATTAGG	419
OY	2148	TACCTTATTAACAGGAAATGCTAAGGTACCTTCTCTGTGAACAATCTGCAATGCTTAAA	2207
Db	420	TACCTTATTAACAGGAAATGCTAAGGTACCTTCTCTGTGAACAATCTGCAATGCTTAAA	479
OY	2208	TCGCCCTTAAAGAGCCCATTTTCTTAGCTGCTGAATCAGTGCCTTTCACCTTCTTCAGAG	2267
Db	480	TCGCCCTTAAAGAGCCCATTTTCTTAGCTGCTGAATCAGTGCCTTTCACCTTCTTCAGAG	539
OY	2268	AAGCAGGGATGTAACCTACCCCGGCAGGTAGGTAGATGTGGGTGGTGCATGTTAATTCC	2327
Db	540	AAGCAGGGATGTAACCTACCCCGGCAGGTAGGTAGATGTGGGTGGTGCATGTTAATTCC	599
OY	2328	CTTAGAAGTTCCAAGCCCTGTTTCTCTGCGTA AAGGTGTATGTCCAGTTCAGAGATGTGT	2387
Db	600	CTTAGAAGTTCCAAGCCCTGTTTCTCTGCGTA AAGGTGTATGTCCAGTTCAGAGATGTGT	659
OY	2388	ATAATGAGCATGGCTGTGTTAAGATCAGGAGGCCACCTGGATTATAGTATAGCCCTTCC	2447
Db	660	ATAATGAGCATGGCTGTGTTAAGATCAGGAGGCCACCTGGATTATAGTATAGCCCTTCC	719
OY	2448	TCCACTCCCAACCAGACTTGCTCATTTTTCGAGTTTAACTAGACTFACACTTATTT	2503
Db	720	TCCACTCCCAACCAGACTTGCTCATTTTTCGAGTTTAACTAGACTFACACTTATTT	775

RESULT 3
US-09-328-111-850/c
; Sequence 850, Application US/09328111

```

; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.

```

APPLICANT: Astle, Jon H.
APPLICANT: Burgess, Christopher C.
APPLICANT: Bushnell, Steven E.
APPLICANT: Carroll III, Eddie
APPLICANT: Catino, Theodore J.
APPLICANT: Dertl, Adnan
APPLICANT: Ford, Donna M.
APPLICANT: Lewis, Marcia E.
APPLICANT: Monahan, John E.
APPLICANT: Schlegel, Robert
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
FILE OF INVENTION: PRODUCTS
FILE REFERENCE: CCD-257 (US)
CURRENT APPLICATION NUMBER: US/09/328,111
CURRENT FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: US 60/088,801
EARLIER FILING DATE: 1998-06-10
NUMBER OF SEQ ID NOS: 850
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 850
LENGTH: 636
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(636)
OTHER INFORMATION: n = A,T,C or G
US-09-328-111-850

Query Match 14.4%; Score 370; DB 4; Length 636;
Best Local Similarity 100.0%; Pred. No. 3.4e-155;
Matches 370; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 756 CTGGGGCTACGCGCGCTGCTGCTGTTAATCACTACTCACTCAGAATGAAGTAGCT 815
DB 370 CTGGGGCTACGCGCGCTGCTGCTGTTAATCACTACTCACTCAGAATGAAGTAGCT 311
QY 816 GCAGCACTGGCTGAGGCTGAGTGCAGTGTCCGCTTGAAGGCGAGTCAGAAGATGAC 875
DB 310 GCAGCACTGGCTGAGGCTGAGTGCAGTGTCCGCTTGAAGGCGAGTCAGAAGATGAC 251
QY 876 TTCTGGTGGTATGACCCGCTGTGTGAACATGGATGGGTGGCAGGCCCAACATGATCCTG 935
DB 250 TTCTGGTGGTATGACCCGCTGTGTGAACATGGATGGGTGGCAGGCCCAACATGATCCTG 191
QY 936 GATGATGGGGAGACTTAACCCACTGGGTTTATAGAAGTATCCAAAGCTGTTAAGAG 995
DB 190 GATGATGGGGAGACTTAACCCACTGGGTTTATAGAAGTATCCAAAGCTGTTAAGAG 131
QY 996 ATCCGAGGCATGTGAGAGAGAGCGTGAAGTGTGTTCACAGGCTGTATCAGCTCTCCAAA 1055
DB 130 ATCCGAGGCATGTGAGAGAGAGCGTGAAGTGTGTTCACAGGCTGTATCAGCTCTCCAAA 71
QY 1056 GCTGGAAGCTGTGTGTTCCGGCCATGAAGTCAATGATTCGTTTACCAACAGAGATT 1115
DB 70 GCTGGAAGCTGTGTGTTCCGGCCATGAAGTCAATGATTCGTTTACCAACAGAGATT 11
QY 1116 GATAACTTGT 1125
DB 10 GATAACTTGT 1

RESULT 4
US-09-404-879A-40/C
Sequence 40, Application US/09404879A
Patent No. 6468546
GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer L.
APPLICANT: King, Gordon E.
APPLICANT: Algate, Paul A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
FILE REFERENCE: 210121.462C2

CURRENT APPLICATION NUMBER: US/09/404,879A
CURRENT FILING DATE: 1999-09-24
NUMBER OF SEQ ID NOS: 393
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 40
LENGTH: 292
TYPE: DNA
ORGANISM: Homo sapien
US-09-404-879A-40

Query Match 9.2%; Score 237; DB 4; Length 292;
Best Local Similarity 100.0%; Pred. No. 4.8e-96;
Matches 237; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2267 GAAGCAGGATGCTACCTACCCGCGAGGTAGATGTGGGTGTCATGTTAATTC 2326
DB 276 GAAGCAGGATGCTACCTACCCGCGAGGTAGATGTGGGTGTCATGTTAATTC 217
QY 2327 CCTTAGAGTCCAGCCCTGTTCTCGCTAAAGGTGATGTCCAGTTCAGAGATGTG 2386
DB 216 CCTTAGAGTCCAGCCCTGTTCTCGCTAAAGGTGATGTCCAGTTCAGAGATGTG 157
QY 2387 TATAATGAGCATGGCTGTTAAGATCAGAGGCCCACTGGATTATATAGCCCTTC 2446
DB 156 TATAATGAGCATGGCTGTTAAGATCAGAGGCCCACTGGATTATATAGCCCTTC 97
QY 2447 CTCGACTCCACAGACTGCTCATTTTTCGAGTTTAACTAGACTACACTCTATT 2503
DB 96 CTCGACTCCACAGACTGCTCATTTTTCGAGTTTAACTAGACTACACTCTATT 40

RESULT 5
US-08-236-427-12
Sequence 12, Application US/08236427
Patent No. 5541310
GENERAL INFORMATION:
APPLICANT: Ward, Eric
APPLICANT: Volrath, Sandra
APPLICANT: Kolzumi, Sain-Ichi
APPLICANT: Tada, Sachio
APPLICANT: Mori, Ichiro
APPLICANT: Iwasaki, Genji
TITLE OF INVENTION: Herbicide Resistant Plants
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/236,427
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/061,644
FILING DATE: 13-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Sprull, W. Murray
REGISTRATION NUMBER: 32,943
REFERENCE/DOCKET NUMBER: S-19558/P1/CGC 1706/CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8615
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 787 base pairs

TYPE: nucleic acid.
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Arabidopsis thaliana
IMMEDIATE SOURCE:
CLONE: pIGPdat.2
US-08-236-427-12

Query Match 1.0%; Score 26; DB 1; Length 787;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2538 TTGTCACAAAAA 2563
DB 761 TTGTCACAAAAA 786

RESULT 6

US-09-328-111-798/c
Sequence 798, Application US/09328111
Patent No. 6262333
GENERAL INFORMATION:
APPLICANT: Endege, Wilson O.
APPLICANT: Steinmann, Kathleen E.
APPLICANT: Astle, Jon H.
APPLICANT: Burgess, Christopher C.
APPLICANT: Bushnell, Steven E.
APPLICANT: Carroll III, Eddie
APPLICANT: Catino, Theodore J.
APPLICANT: Dertl, Adnan
APPLICANT: Ford, Donna M.
APPLICANT: Lewis, Marcia E.
APPLICANT: Monahan, John E.
APPLICANT: Schlegel, Robert
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
TITLE OF INVENTION: PRODUCTS
FILE REFERENCE: CCD-257 (US)
CURRENT APPLICATION NUMBER: US/09/328,111
CURRENT FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: US 60/088,801
EARLIER FILING DATE: 1998-06-10
NUMBER OF SEQ ID NOS: 850
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 798
LENGTH: 632
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc-feature
LOCATION: (1)...(632)
OTHER INFORMATION: n - A,T,C or G
US-09-328-111-798

Query Match 1.0%; Score 25; DB 4; Length 632;
Best Local Similarity 100.0%; Pred. No. 0.085;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2539 TGTGCAAAAAA 2563
DB 40 TGTGCAAAAAA 16

RESULT 7

US-09-227-357-66/c
Sequence 66, Application US/09227357
Patent No. 6342581
GENERAL INFORMATION:
APPLICANT: Fischer et al.
TITLE OF INVENTION: 123 Human Secreted Proteins

FILE REFERENCE: P2010P1
CURRENT APPLICATION NUMBER: US/09/227,357
CURRENT FILING DATE: 1999-01-08
EARLIER APPLICATION NUMBER: PCT/US98/13684
EARLIER FILING DATE: 1998-07-07
EARLIER APPLICATION NUMBER: 60/051,926
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,793
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,925
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,929
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,803
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,932
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,931
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,930
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,918
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,920
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,733
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,795
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,919
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,928
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/055,722
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,723
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,948
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,949
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,953
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,950
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,947
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,964
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/056,360
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,684
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,984
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,954
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/058,785
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,664
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,660
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,661
EARLIER FILING DATE: 1997-09-12
NUMBER OF SEQ ID NOS: 672
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 66
LENGTH: 685

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (7)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-227-357-66

Query Match 1.0%; Score 25; DB 4; Length 685;
Best Local Similarity 100.0%; Pred. No. 0.085;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2539 TGTCCAAAAA 2563
|||||
Db 78 TGTCCAAAAA 54

RESULT 8
US-09-345-882-3/c
; Sequence 3, Application US/09345882
; Patent No. 6399373
; GENERAL INFORMATION:
; APPLICANT: Bouquellet, Lydie
; TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)
; TITLE OF INVENTION: AND POLYMORPHIC MARKERS ASSOCIATED WITH SAID NUCLEIC ACID.
; FILE REFERENCE: GENSET.031A
; CURRENT APPLICATION NUMBER: US/09/345,882
; CURRENT FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: US 60/091,315
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/111,909
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: Patent.pm
; SEQ ID NO 3
; LENGTH: 999
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-345-882-3

Query Match 1.0%; Score 25; DB 4; Length 999;
Best Local Similarity 100.0%; Pred. No. 0.083;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2539 TGTCCAAAAA 2563
|||||
Db 208 TGTCCAAAAA 184

RESULT 9
US-09-262-773-209
; Sequence 209, Application US/09262773
; Patent No. 6225451
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis G.
; APPLICANT: Ding, Wei
; APPLICANT: Wagner, Susanne
; APPLICANT: Hess, Mark A.
; TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
; TITLE OF INVENTION: SUSCEPTIBILITY GENE CHD1
; FILE REFERENCE: Myriad 3
; CURRENT APPLICATION NUMBER: US/09/262,773
; CURRENT FILING DATE: 1999-03-04
; NUMBER OF SEQ ID NOS: 210
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 209
; LENGTH: 2933
; TYPE: DNA
; ORGANISM: human
US-09-262-773-209

Query Match 1.0%; Score 25; DB 4; Length 2933;
Best Local Similarity 100.0%; Pred. No. 0.079;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2539 TGTCCAAAAA 2563
|||||
Db 920 TGTCCAAAAA 944

RESULT 10
US-09-262-773-210
; Sequence 210, Application US/09262773
; Patent No. 6225451
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis G.
; APPLICANT: Ding, Wei
; APPLICANT: Wagner, Susanne
; APPLICANT: Hess, Mark A.
; TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
; TITLE OF INVENTION: SUSCEPTIBILITY GENE CHD1
; FILE REFERENCE: Myriad 3
; CURRENT APPLICATION NUMBER: US/09/262,773
; CURRENT FILING DATE: 1999-03-04
; NUMBER OF SEQ ID NOS: 210
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 210
; LENGTH: 23071
; TYPE: DNA
; ORGANISM: human
US-09-262-773-210

Query Match 1.0%; Score 25; DB 4; Length 23071;
Best Local Similarity 100.0%; Pred. No. 0.073;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2539 TGTCCAAAAA 2563
|||||
Db 920 TGTCCAAAAA 944

RESULT 11
US-09-345-882-1/c
; Sequence 1, Application US/09345882
; Patent No. 6399373
; GENERAL INFORMATION:
; APPLICANT: Bouquellet, Lydie
; TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)
; TITLE OF INVENTION: AND POLYMORPHIC MARKERS ASSOCIATED WITH SAID NUCLEIC ACID.
; FILE REFERENCE: GENSET.031A
; CURRENT APPLICATION NUMBER: US/09/345,882
; CURRENT FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: US 60/091,315
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/111,909
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 162450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 72794
; OTHER INFORMATION: 5-124-273 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 88073
; OTHER INFORMATION: 5-127-261 : polymorphic base A or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 90842
; OTHER INFORMATION: 99-1437-325 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele

LOCATION: 93714
OTHER INFORMATION: 5-128-60 : polymorphic base deletion of GT
FEATURE:
NAME/KEY: allele
LOCATION: 97122
OTHER INFORMATION: 99-1442-224 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 97152
OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T
FEATURE:
NAME/KEY: allele
LOCATION: 99098
OTHER INFORMATION: 5-130-257 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 99117
OTHER INFORMATION: 5-130-276 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 103806
OTHER INFORMATION: 5-131-395 : polymorphic base A or T
FEATURE:
NAME/KEY: allele
LOCATION: 106940
OTHER INFORMATION: 5-133-375 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108106
OTHER INFORMATION: 5-135-155 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108149
OTHER INFORMATION: 5-135-198 : polymorphic base insertion of GTTT
FEATURE:
NAME/KEY: allele
LOCATION: 108308
OTHER INFORMATION: 5-135-357 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 108471
OTHER INFORMATION: 5-136-174 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134134
OTHER INFORMATION: 5-140-120 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134362
OTHER INFORMATION: 5-140-348 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 134374
OTHER INFORMATION: 5-140-361 : polymorphic base insertion of CA
FEATURE:
NAME/KEY: allele
LOCATION: 146328
OTHER INFORMATION: 5-143-84 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 146345
OTHER INFORMATION: 5-143-101 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 150329
OTHER INFORMATION: 5-145-24 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 160031
OTHER INFORMATION: 5-148-352 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 72771..72817

OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30
FEATURE:
NAME/KEY: allele
LOCATION: 72771..72817
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
FEATURE:
NAME/KEY: allele
LOCATION: 88050..88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
FEATURE:
NAME/KEY: allele
LOCATION: 88050..88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37

FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
FEATURE:

Query Match 1.0%; Score 25; DB 4; Length 162450;
Best Local Similarity 100.0%; Pred. No. 0.067;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2539 TGTCCAAAAA 2563
DB 161659 TGTCCAAAAA 161635

RESULT 12

US-08-776-944-9/c
Sequence 9, Application US/08776944
Patent No. 5925520

GENERAL INFORMATION:

APPLICANT: TULLY, GILLIAN
APPLICANT: SULLIVAN, KEVIN
TITLE OF INVENTION: NUCLEOTIDE SEQUENCING METHOD
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/776,944
FILING DATE: 14-FEB-1997
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9417211.1
FILING DATE: 25-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: CRAWFORD, ARTHUR R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 1498-92
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4006
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:

LENGTH: 56 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
ORGANELLE: Mitochondrion
US-08-776-944-9

Query Match 0.9%; Score 23; DB 2; Length 56;
Best Local Similarity 100.0%; Pred. No. 0.73;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2541 TCCAAAAA 2563
DB 32 TCCAAAAA 10

RESULT 13

US-07-920-281C-25
Sequence 25, Application US/07920281C
Patent No. 5739026

GENERAL INFORMATION:

APPLICANT: Garoff, Henrik
APPLICANT: Liljestrom, Peter
TITLE OF INVENTION: DNA Expression Systems Based on
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:

ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: P.O. Box 747
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22040-0747

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/920,281C
FILING DATE: 13-AUG-1992
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Murphy, Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 828-103P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-241-1300
TELEFAX: 703-241-2848
TELEX: 248345

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:
LENGTH: 80 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: RNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Semliki Forest Virus

FEATURE:

NAME/KEY: 1..80
LOCATION: 1..80
OTHER INFORMATION: /label= terminator
OTHER INFORMATION: /note= "3' terminal sequence of cDNA expression
OTHER INFORMATION: vector complementary to alphavirus genomic RNA"
US-07-920-281C-25

Query Match 0.9%; Score 23; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.72;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2541 TCCAAAAAAAAAAAAAAAAAAAAA 2563
|||||
Db 3 TCCAAAAAAAAAAAAAAAAAAAAA 25

RESULT 14

US-08-466-277-25

; Sequence 25, Application US/08466277
; Patent No. 6190666

GENERAL INFORMATION:

APPLICANT: Garoff, Henrik

Liljestrom, Peter

TITLE OF INVENTION: DNA Expression Systems Based on

Alphaviruses

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:

ADDRESSEE: Birch, Stewart, Kolasch & Birch

STREET: P.O. Box 747

CITY: Falls Church

STATE: Virginia

COUNTRY: USA

ZIP: 22040-0747

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/466,277

FILING DATE: 06-Jun-1995

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/920,281

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Murphy Jr., Gerald M.

REGISTRATION NUMBER: 28,977

REFERENCE/DOCKET NUMBER: 828-103P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-241-1300

TELEFAX: 703-241-2848

TELEX: 248345

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:

LENGTH: 80 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: RNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Semliki Forest Virus

FEATURE:

NAME/KEY: -

LOCATION: 1..80

OTHER INFORMATION: /label- terminator

/note- "3' terminal sequence of cDNA expression

vector complementary to alphavirus genomic RNA"

SEQUENCE DESCRIPTION: SEQ ID NO: 25:

US-08-466-277-25

Query Match

Best Local Similarity 100.0%; Score 23; DB 4; Length 80;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2541 TCCAAAAAAAAAAAAAAAAAAAAA 2563

|||||

Db 3 TCCAAAAAAAAAAAAAAAAAAAAA 25

RESULT 15

US-09-385-982-187/c

; Sequence 187, Application US/09385982
; Patent No. 6262334

GENERAL INFORMATION:

APPLICANT: ENDEGE, WILSON O., ET AL.

TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION

FILE REFERENCE: CCDNA-260XX

CURRENT APPLICATION NUMBER: US/09/385,982

CURRENT FILING DATE: 1999-08-30

EARLIER APPLICATION NUMBER: 09/328,111

EARLIER FILING DATE: 1999-06-08

EARLIER APPLICATION NUMBER: 60/117,393

EARLIER FILING DATE: 1999-01-27

EARLIER APPLICATION NUMBER: 60/098,639

EARLIER FILING DATE: 1998-08-31

NUMBER OF SEQ ID NOS: 544

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 187

LENGTH: 256

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1)..(256)

OTHER INFORMATION: n - A,T,C or G

US-09-385-982-187

Query Match

Best Local Similarity 100.0%; Score 23; DB 4; Length 256;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2541 TCCAAAAAAAAAAAAAAAAAAAAA 2563

|||||

Db 3 TCCAAAAAAAAAAAAAAAAAAAAA 11

Search completed: April 20, 2003, 21:12:54
Job time : 185 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2003, 21:12:59 ; Search time 268 Seconds

(without alignments)
9607.061 Million cell updates/sec

Title: US-09-782-051-1
Perfect score: 2563
Sequence: 1 ggcgcggcgagtcgcgagct.....aaaaaaaaaaaaaaaaa 2563

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 639749 seqs, 502280978 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1279498

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications_NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2563	100.0	2563	10	US-09-782-051-1 Sequence 1, Appl1
2	570	22.2	721	10	US-09-925-300-439 Sequence 439, App
3	370	14.4	636	10	US-09-879-536-850 Sequence 850, App
4	330	12.9	376	10	US-09-880-107-1031 Sequence 1031, Ap
5	237	9.2	292	10	US-09-884-441-40 Sequence 40, Appl
6	224	8.7	553	10	US-09-867-701-2023 Sequence 2023, Ap
7	190	7.4	1877	12	US-10-021-121-1 Sequence 1, Appl1
8	168	6.6	384	10	US-09-860-107-346 Sequence 346, App
9	168	6.6	140	10	US-09-783-590-278 Sequence 278, App
10	45	1.8	863	9	US-10-114-893-33 Sequence 33, Appl
11	26	1.0	987	10	US-09-770-445-261 Sequence 261, App
12	26	1.0	1165	10	US-09-847-519A-1 Sequence 1, Appl1
13	26	1.0	2000	9	US-09-938-842A-3304 Sequence 3304, Ap
14	26	1.0	2003	10	US-09-887-576-15 Sequence 15, Appl
15	25	1.0	101	10	US-09-783-590-4696 Sequence 4696, Ap
16	25	1.0	187	9	US-09-736-457-1474 Sequence 1474, Ap
17	25	1.0	187	9	US-09-902-941-1474 Sequence 1474, Ap
18	25	1.0	187	9	US-09-849-626-1474 Sequence 1474, Ap
19	25	1.0	187	9	US-10-017-754-1474 Sequence 1474, Ap

C	20	25	1.0	294	10	US-09-867-701-10512	Sequence 10512, A
C	21	25	1.0	353	10	US-09-925-301-43	Sequence 43, Appl
C	22	25	1.0	632	10	US-09-879-536-798	Sequence 798, App
C	23	25	1.0	667	10	US-09-770-149-438	Sequence 438, App
C	24	25	1.0	685	9	US-09-983-802-66	Sequence 66, Appl
C	25	25	1.0	2000	9	US-09-938-842A-5235	Sequence 5235, Ap
C	26	25	1.0	2683	12	US-10-012-452-1	Sequence 1, Appl1
C	27	25	1.0	3044	9	US-09-125-540-98	Sequence 98, Appl
C	28	25	1.0	3044	10	US-09-764-870-98	Sequence 98, Appl
C	29	25	1.0	3150	9	US-10-125-540-622	Sequence 622, App
C	30	25	1.0	3150	10	US-09-764-870-622	Sequence 622, App
C	31	25	1.0	74962	9	US-10-274-974-3	Sequence 3, Appl1
C	32	24	0.9	167	10	US-09-867-701-8875	Sequence 8875, Ap
C	33	24	0.9	188	9	US-09-768-827-7	Sequence 7, Appl1
C	34	24	0.9	193	10	US-09-983-965-3122	Sequence 3122, Ap
C	35	24	0.9	208	10	US-09-867-701-8988	Sequence 8988, Ap
C	36	24	0.9	216	10	US-09-867-701-9355	Sequence 9355, Ap
C	37	24	0.9	240	9	US-10-015-219-855	Sequence 855, App
C	38	24	0.9	240	10	US-09-777-564-855	Sequence 855, App
C	39	24	0.9	275	10	US-09-834-975-619	Sequence 619, App
C	40	24	0.9	300	10	US-09-864-864-126	Sequence 126, App
C	41	24	0.9	303	10	US-09-864-864-161	Sequence 161, App
C	42	24	0.9	428	9	US-09-525-978B-74	Sequence 74, Appl
C	43	24	0.9	523	10	US-09-834-975-947	Sequence 947, App
C	44	24	0.9	707	9	US-10-001-887-12	Sequence 12, Appl
C	45	24	0.9	743	10	US-09-867-550-1185	Sequence 1185, Ap

ALIGNMENTS

RESULT 1
US-09-782-051-1
Sequence 1, Application US/09782051
Patent No. US20020035078A1
GENERAL INFORMATION:
APPLICANT: Hart, Derek N J
TITLE OF INVENTION: Enzyme having S-adenosyl-L-homocysteine hydrolase
TITLE OF INVENTION: (AHCY) type activity
FILE REFERENCE: 24305 MRB
CURRENT APPLICATION NUMBER: US/09/782,051
CURRENT FILING DATE: 2001-02-14
PRIOR APPLICATION NUMBER: PCT/NZ97/00133
PRIOR FILING DATE: 1997-10-06
PRIOR APPLICATION NUMBER: NZ 299507
PRIOR FILING DATE: 1996-10-04
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 2563
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (3)..(1847)
OTHER INFORMATION: Open reading frame extends without a stop codon
OTHER INFORMATION: for the full 5' nucleotide sequence. The
OTHER INFORMATION: Initiation codon has yet to be identified.
US-09-782-051-1

Query Match 100.0%; Score 2563; DB 10; Length 2563;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GGCGGGGCGAGTCGAGCTCGAGCTGCTTCTGTTCTTGTGGCCACCGTCGCT	60
DB	1	GGCGGGGCGAGTCGAGCTCGAGCTGCTTCTGTTCTTGTGGCCACCGTCGCT	60
QY	61	GTCGCGCTGCTTGGCTGCGGAACAGACAGAGGCTGGCCACAGCACCCTCAGAGCCGA	120
DB	61	GTCGCGCTGCTTGGCTGCGGAACAGACAGAGGCTGGCCACAGCACCCTCAGAGCCGA	120
QY	121	CGCAGCTCGACGACGAGGCGCCGACAGAGGCTGGCGCATCGCTGTCGAGAGCGCGCGC	180

Db 121 CGCAGCTCGACCGAGGGCGCGCAGAGGGGTGGCGCATCGCGTGTGGAGGGCGCGCGC 180
QY 181 GGGCAGCGCGCGCGCGCGCAGAGGGGGAAGAGCGCGGGCGCGGGTTCAGCCGCTGGCC 240
Db 181 GGGCAGCGCGCGCGCGCGCAGAGGGGGAAGAGCGCGGGCGCGGGTTCAGCCGCTGGCC 240
QY 241 GGGCGCGCGGGGAATGTCGATGCCCTGACGCGCATGCCGCTGCCCGGGGTGGGAGAGC 300
Db 241 GGGCGCGCGGGGAATGTCGATGCCCTGACGCGCATGCCGCTGCCCGGGGTGGGAGAGC 300
QY 301 TGAAGCAGGCCAAGAGATCGAGGACGCCGAGAAGTACTCCTTCATGCGCCACCGTCACCA 360
Db 301 TGAAGCAGGCCAAGAGATCGAGGACGCCGAGAAGTACTCCTTCATGCGCCACCGTCACCA 360
QY 361 AGCGCGCCCAAGAACCAATCCAGTTGCTGATGACATGCAAGAGTTCCACCAATTCGCCA 420
Db 361 AGCGCGCCCAAGAACCAATCCAGTTGCTGATGACATGCAAGAGTTCCACCAATTCGCCA 420
QY 421 CCAAAACTGGCGGAAGATCTTGTCTCGCTCGATCTCACAGTCCCTCCACTGACAGCTACA 480
Db 421 CCAAACTGGCGGAAGATCTTGTCTCGCTCGATCTCACAGTCCCTCCACTGACAGCTACA 480
QY 481 GTTCAGCTGCATCCTACACAGATAGCTGATGATGAGGTTTCTCCCGAGAGAAGCAGC 540
Db 481 GTTCAGCTGCATCCTACACAGATAGCTGATGATGAGGTTTCTCCCGAGAGAAGCAGC 540
QY 541 AAACCACTCCAAAGGCGACGACCAATTTCTGTGGAAGACATCAAGCAGAGAAATTG 600
Db 541 AAACCACTCCAAAGGCGACGACCAATTTCTGTGGAAGACATCAAGCAGAGAAATTG 600
QY 601 GACGCCGGGAGATTGAGATTGCAGAGCAGACATGCTGCTGTGATTTCACTCAGGAAC 660
Db 601 GACGCCGGGAGATTGAGATTGCAGAGCAGACATGCTGCTGTGATTTCACTCAGGAAC 660
QY 661 GTGCTCAGGGGGAAGAGCCCTTGGCTGGTGTAAATAGTGGGCTGTACACATCACAG 720
Db 661 GTGCTCAGGGGGAAGAGCCCTTGGCTGGTGTAAATAGTGGGCTGTACACATCACAG 720
QY 721 CCCAGACAGCGTGTGATTGAGACACTGTCCTGGGCTCAGTCCGCTGCTG 780
Db 721 CCCAGACAGCGTGTGATTGAGACACTGTCCTGGGCTCAGTCCGCTGCTG 780
QY 781 CTGTAACTACTCACTCAGAATGAATGAGTGCAGCACTGGCTGAGGCTGAGT 840
Db 781 CTGTAACTACTCACTCAGAATGAATGAGTGCAGCACTGGCTGAGGCTGAGT 840
QY 841 CAGTGTCCGTTGGAAGGGCGAGTCAGAGATGACTTCTGGTGTGATTTGACCGCTG 900
Db 841 CAGTGTCCGTTGGAAGGGCGAGTCAGAGATGACTTCTGGTGTGATTTGACCGCTG 900
QY 901 TGAACATGATGGGTGGCAGGCCCAACATGATCTGATGATGGGGAGACTTAACCCACT 960
Db 901 TGAACATGATGGGTGGCAGGCCCAACATGATCTGATGATGGGGAGACTTAACCCACT 960
QY 961 GGGTTTAAAGAGTATCCAAACGTGTTAAAGATCCGAGGCAATGGAAGAGCG 1020
Db 961 GGGTTTAAAGAGTATCCAAACGTGTTAAAGATCCGAGGCAATGGAAGAGCG 1020
QY 1021 TGACTGCTGTCACAGGCTGATCAGCTCTCCAAAGCTGGGAAGCTGTGTCCGCCA 1080
Db 1021 TGACTGCTGTCACAGGCTGATCAGCTCTCCAAAGCTGGGAAGCTGTGTCCGCCA 1080
QY 1081 TGAACGTCAATGATTTCTGTACCAACAGAAATTGATACTGTACTGCTCCGAGAA 1140
Db 1081 TGAACGTCAATGATTTCTGTACCAACAGAAATTGATACTGTACTGCTCCGAGAA 1140
QY 1141 CCATTGATGAGGCTGAAGAGGACCAAGATGATGTTGGTGGAAACAAGTGTG 1200
Db 1141 CCATTGATGAGGCTGAAGAGGACCAAGATGATGTTGGTGGAAACAAGTGTG 1200
QY 1201 TGTGTGCTATGATGAGTAGGCAAGGCTGCTGCTCAAGCTCTTGAGCAA 1260
Db 1201 TGTGTGCTATGATGAGTAGGCAAGGCTGCTGCTCTCAAGCTCTTGAGCAA 1260

Db 1201 TGTGTGCTATGATGAGTAGGCAAGGCTGCTGCTCTCTCAAGCTCTTGAGCAA 1260
QY 1261 TTGCTACATTACCGAAATCGAACCCCATCTGTGCTCTGACAGGCTGCATGATGGGTCA 1320
Db 1261 TTGCTACATTACCGAAATCGAACCCCATCTGTGCTCTGACAGGCTGCATGATGGGTCA 1320
QY 1321 GGGTGTAAAGCTAAATGAAGTCATCCGGCAAGTCGATGCTAATTAATTGACAGGAA 1380
Db 1321 GGGTGTAAAGCTAAATGAAGTCATCCGGCAAGTCGATGCTAATTAATTGACAGGAA 1380
QY 1381 ATAAGATGTAGTACACGGGAGCAGCTTGATGCGATGAAGAAACAGTTGATCGATGCA 1440
Db 1381 ATAAGATGTAGTACACGGGAGCAGCTTGATGCGATGAAGAAACAGTTGATCGATGCA 1440
QY 1441 ATATGGCCACTCCAAACAGAAATCGATGTGACAGCCTCCGCACTCCGAGCTGACGT 1500
Db 1441 ATATGGCCACTCCAAACAGAAATCGATGTGACAGCCTCCGCACTCCGAGCTGACGT 1500
QY 1501 GGGAGCGAGTACGTTCTCAGGTGAGCAGTGTCACTGCGCCAGATGGCAACGAGTTGCC 1560
Db 1501 GGGAGCGAGTACGTTCTCAGGTGAGCAGTGTCACTGCGCCAGATGGCAACGAGTTGCC 1560
QY 1561 TCCTGGCAGAGGGTCTCTACTCAATTGAGCTGCTCCACAGTTCCCACTTTGTTCTGT 1620
Db 1561 TCCTGGCAGAGGGTCTCTACTCAATTGAGCTGCTCCACAGTTCCCACTTTGTTCTGT 1620
QY 1621 CCATCACAGCCACAACAGGCTTTGGCAGCTGATGAGACTCTATATGACCCGAGGGC 1680
Db 1621 CCATCACAGCCACAACAGGCTTTGGCAGCTGATGAGACTCTATATGACCCGAGGGC 1680
QY 1681 GATACAAGCAGAGTGTACTTGTCTTCTTAAGAAATGATGAATACGTTGCCAGCTTC 1740
Db 1681 GATACAAGCAGAGTGTACTTGTCTTCTTAAGAAATGATGAATACGTTGCCAGCTTC 1740
QY 1741 ATCTGCCATCATTTGATGCCCACTTACAGAGCTGACAGATGACCAAGCAAAATCTGG 1800
Db 1741 ATCTGCCATCATTTGATGCCCACTTACAGAGCTGACAGATGACCAAGCAAAATCTGG 1800
QY 1801 GACTCAACAAAATGGGCCATTCAACCTAATTATACAGATCTAATGAGCATACTAC 1860
Db 1801 GACTCAACAAAATGGGCCATTCAACCTAATTATACAGATCTAATGAGCATACTAC 1860
QY 1861 CAAGGACAGTCCACCTGAACACACACACTCTAAAGAAATATTTTAAGATACTTTAT 1920
Db 1861 CAAGGACAGTCCACCTGAACACACACACTCTAAAGAAATATTTTAAGATACTTTAT 1920
QY 1921 TTTCTTCTACTCCTTCCCTCTGATTTTTCCTAATTAATTCACTCTGTTTTCATC 1980
Db 1921 TTTCTTCTACTCCTTCCCTCTGATTTTTCCTAATTAATTCACTCTGTTTTCATC 1980
QY 1981 TCATTATCCAAAGTTCTGACAGACACACAGAACTTCCTCATGGCTCTTAGATGAATA 2040
Db 1981 TCATTATCCAAAGTTCTGACAGACACACAGAACTTCCTCATGGCTCTTAGATGAATA 2040
QY 2041 GAAGTCAAGGTCCTCACTCTAGTCACTAAGAAAGATTTTACTCCCCAGCCAGAAA 2100
Db 2041 GAAGTCAAGGTCCTCACTCTAGTCACTAAGAAAGATTTTACTCCCCAGCCAGAAA 2100
QY 2101 GGTGATTTCTCTTACCATTTCTGGGACTTTAGTCTTAATTAGGTACTTAATAACA 2160
Db 2101 GGTGATTTCTCTTACCATTTCTGGGACTTTAGTCTTAATTAGGTACTTAATAACA 2160
QY 2161 GGAATGCTAAGTACCTCTCTGTGAACAATGCAATGTCTAATCGCCTTAAGAA 2220
Db 2161 GGAATGCTAAGTACCTCTCTGTGAACAATGCAATGTCTAATCGCCTTAAGAA 2220
QY 2221 GCCCATTTCTTACGCTGTAATCAGTGTCTTCACTTCTCAGAGAGAGGAGTGGT 2280
Db 2221 GCCCATTTCTTACGCTGTAATCAGTGTCTTCACTTCTCAGAGAGAGGAGTGGT 2280
QY 2281 ACCTACCCGGCAGGATGATGATGTTGGTGTGATGTTAATTCCCTTAGAAGTTCCA 2340
Db 2281 ACCTACCCGGCAGGATGATGATGTTGGTGTGATGTTAATTCCCTTAGAAGTTCCA 2340

QY 2341 AGCCCTGTTCTCCGTAAGGTGTATGTCCAGTTCAGAGATGTGTATATGAGCATGG 2400
DB 2341 AGCCCTGTTCTCCGTAAGGTGTATGTCCAGTTCAGAGATGTGTATATGAGCATGG 2400
QY 2401 CTGTGTAAGATCAGAGAGCCCACTTGATTTATAGTATAGCCCTTCTCCACTCCACCA 2460
DB 2401 CTGTGTAAGATCAGAGAGCCCACTTGATTTATAGTATAGCCCTTCTCCACTCCACCA 2460
QY 2461 GACTTGCTCATTTTTCGATTTTAACTAGACTACACTATTTGAGTTTATTTTGTCC 2520
DB 2461 GACTTGCTCATTTTTCGATTTTAACTAGACTACACTATTTGAGTTTATTTTGTCC 2520
QY 2521 TCTAGGATTTATTTCTGTGTCCAAAAA 2563
DB 2521 TCTAGGATTTATTTCTGTGTCCAAAAA 2563

RESULT 2

US-09-925-300-439
; Sequence 439, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 439
; LENGTH: 721
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (688)
; OTHER INFORMATION: -n equals a,t,g, or c
US-09-925-300-439

Query Match 22.2%; Score 570; DB 10; Length 721;
Best Local Similarity 99.8%; Pred. No. 1.5e-292;
Matches 620; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 53 CCGTCGCTGTCGGCTGCTTGGGCTGCCGAACAGAGCGGTGGCCACAGACACTCA 112
DB 56 CCGTCGCTGTCGGCTGCTTGGGCTGCCGAACAGAGCGGTGGCCACAGACACTCA 115
QY 113 GAAGCCGACGACGTCGACGAGGGGCGGCGGAGAGGCTGGCGGATCGCTGTCGAGGG 172
DB 116 GAAGCCGACGACGTCGACGAGGGGCGGCGGAGAGGCTGGCGGATCGCTGTCGAGGG 175
QY 173 CGCCGCGGCGGAGGGGCGGCGGCGGAGAGGGGAGAGGGGCGGCGGCGGCTCAGC 232
DB 176 CGCCGCGGCGGAGGGGCGGCGGCGGAGAGGGGAGAGGGGCGGCGGCGGCTCAGC 235
QY 233 CGGCGGCGGCGGCGGCGGCGGAGATGTGATGCTGACGCGATGCCGCTGCCGCGGCTCGG 292
DB 236 CGCTGGCGGCGGCGGCGGCGGAGATGTGATGCTGACGCGATGCCGCTGCCGCGGCTCGG 295
QY 293 GGAGGAGCTGAAGAGCCCAAGAGATGAGAGAGCGGAGAGAGTACTCTTCATGGCCAC 352
DB 296 GGAGGAGCTGAAGAGCCCAAGAGATGAGAGAGCGGAGAGAGTACTCTTCATGGCCAC 355
QY 353 CGTCACCAAGGCGCCCAAGAGCAATCCAGTTGCTGATGACATGACAGAGTTCACCAA 412
DB 356 CGTCACCAAGGCGCCCAAGAGCAATCCAGTTGCTGATGACATGACAGAGTTCACCAA 415

QY 413 ATTCCCAACCAAACTGGCCGAGAGATCTTGTCTCGCTCGATCTCAGACTCTCCACTGA 472
DB 416 ATTCCCAACCAAACTGGCCGAGAGATCTTGTCTCGCTCGATCTCAGACTCTCCACTGA 475
QY 473 CAGCTACAGTTTCACTGCTGATCTCTACACAGATAGCTCTGATGATGAGGTTTCTCCCGAGA 532
DB 476 CAGCTACAGTTTCACTGCTGATCTCTACACAGATAGCTCTGATGATGAGGTTTCTCCCGAGA 535
QY 533 GAAGCAGCAAACTCACTCCAAAGGCGAGAGCAATTTCTGTGTGAAGACATCAAGCAGGC 592
DB 536 GAAGCAGCAAACTCACTCCAAAGGCGAGAGCAATTTCTGTGTGAAGACATCAAGCAGGC 595
QY 593 AGAATTTGACGCGCGGAGATTGAGATTGACAGCAAGACATGCTGCTGATTTCAC 652
DB 596 AGAATTTGACGCGCGGAGATTGAGATTGACAGCAAGACATGCTGCTGATTTCAC 655
QY 653 CAGAAACGTCCTCAGGGGA 673
DB 656 CAGAAACGTCCTCAGGGGA 676

RESULT 3

US-09-879-536-850/c
; Sequence 850, Application US/09879536
; Patent No. US20020144298A1
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steilmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Carino, Theodore J.
; APPLICANT: Dertl, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/879,536
; CURRENT FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/088,801
; PRIOR FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 850
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(636)
; OTHER INFORMATION: n = A,T,C or G
US-09-879-536-850

Query Match 14.4%; Score 370; DB 10; Length 636;
Best Local Similarity 100.0%; Pred. No. 3.6e-186;
Matches 370; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 756 CTGGGCGCTCAGTGCCTGCTGCTTGTATACATCTACTCACTCAGATGAGTAGCT 815
DB 370 CTGGGCGCTCAGTGCCTGCTGCTTGTATACATCTACTCACTCAGATGAGTAGCT 811
QY 816 GCAGCACTGCTGAGGCTGAGTTCAGTGTTCGCTTGAAGGCGAGTCAGAGATGAC 875
DB 310 GCAGCACTGCTGAGGCTGAGTTCAGTGTTCGCTTGAAGGCGAGTCAGAGATGAC 875
QY 876 TTCTGCTGTGTATGACCGCTGTGTGAACATGATGGGTGAGGCGCAACATGATCTG 935
DB 250 TTCTGCTGTGTATGACCGCTGTGTGAACATGATGGGTGAGGCGCAACATGATCTG 931

[illegible]

RESULT 4

```

US-09-880-107-1031/c
; Sequence 1031, Application US/09880107
; Patent NO. US20020142981A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Horne, Darcl T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1031
;
; LENGTH: 376
;
; TYPE: DNA
;
; ORGANISM: Homo sapiens
;
; FEATURE:
;
; OTHER INFORMATION: Genbank Accession NO. US20020142981A1 AA432168
US-09-880-107-1031

```

Query Match	12.9%	Score 330;	DB 10;	Length 376;
Best Local Similarity	100.0%	Pred. No. 6.6e-165;		
Matches 330; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

OY	2174	TACCTTCTCTGTGGAACAATCTGC AATGTCTAAATCGCCTTAAAAAGAGCCCATTTCTTAG	2233
Db	376	TACCTTCTCTGTGGAACAATCTGC AATGTCTAAATCGCCTTAAAAAGAGCCCATTTCTTAG	317
OY	2234	CTGCTGAATCAGTGCTCTTTC ACTTCTTCAGAGAAGCAGGATGGTACC TACCCGGCAG	2293
Db	316	CTGCTGAATCAGTGCTCTTTC ACTTCTTCAGAGAAGCAGGATGGTACC TACCCGGCAG	257
OY	2294	G TAGGTTAGATGTGGGTGGTG CATGTATTTCCTTAGAAGTCCCAAGCCCTGTTCCCT	2353
Db	256	G TAGGTTAGATGTGGGTGGTG CATGTATTTCCTTAGAAGTCCCAAGCCCTGTTCCCT	197
OY	2354	GCGTA AAGGTGGTATGTC CAGTTCAGAGATGTGTATTA TGAGCATGGCTTGTTAAGATCA	2413
Db	196	GCGTA AAGGTGGTATGTC CAGTTCAGAGATGTGTATTA TGAGCATGGCTTGTTAAGATCA	137
OY	2414	GGAGGCCCACCTGGATTATTA GTATATAGCCCTTCCTCCACTCCCAC CAGACTTGCTCATTT	2473
Db	136	GGAGGCCCACCTGGATTATTA GTATATAGCCCTTCCTCCACTCCCAC CAGACTTGCTCATTT	77
OY	2474	TTGAGTTTTTA ACTAGACTAC CTCTATT	2503
Db	76	TTGAGTTTTTA ACTAGACTAC CTCTATT	47

```

RESULT 5
US-09-884-441-40/c
; Sequence 40, Application US/09884441
; Patent NO. US20020119158A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Algate, Paul A.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C7
;
; CURRENT APPLICATION NUMBER: US/09/884,441
;
; CURRENT FILING DATE: 2001-06-18
;
; NUMBER OF SEQ ID NOS: 489
;
; SOFTWARE: FastSeq for Windows Version 3.0
;
; SEQ ID NO 40
;
; LENGTH: 292
;
; TYPE: DNA
;
; ORGANISM: Homo sapien
;
US-09-884-441-40

```

Query Match	9.2%;	Score 237;	DB 10;	Length 292;
Best Local Similarity	100.0%;	Pred. No. 1.9e-115;		

Best Local Similarity 100.0%; Pred. No. 1.9e-115;
Matches 237; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	2267	GAAGCAGGATGGTACTACCCCGCAGGTAGGTAGATGTGGTGCTCATGTTAATTTC	2326
Db	276	GAAGCAGGATGGTACTACCCGCGCAGGTAGGTAGATGTGGTGCTCATGTTAATTTC	217
QY	2327	CCTTAGAGTTCCAAGCCCTGTTTCTGCGTAAGGTGATGTCCAGTTCAGAGATGTG	2386
Db	216	CCTTAGAGTTCCAAGCCCTGTTTCTGCGTAAGGTGATGTCCAGTTCAGAGATGTG	157
QY	2387	TATATAGAGCATGGCTTGTTAAGATCAGGAGGCCCACTTGGATTATAGTATAGCCCCTTC	2446
Db	156	TATATAGAGCATGGCTTGTTAAGATCAGGAGGCCCACTTGGATTATAGTATAGCCCCTTC	97
QY	2447	CTCCACTCCCAACCAGAAGCTGCTCATTTTTTCGAGTTTTTAACTAGACTACACTCTATT	2503
Db	96	CTCCACTCCCAACCAGAAGCTGCTCATTTTTTCGAGTTTTTAACTAGACTACACTCTATT	40

RESULT 6

```

US-09-867-701-2023
; Sequence 2023, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2023
; LENGTH: 553
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(553)
; OTHER INFORMATION: n - A,T,C or G
US-09-867-701-2023

```

Query Match	8.7%;	Score 224;	DB 10;	Length 553;
Best Local Similarity	100.0%;	Pred. No. 1.6e-108;		
Matches 224; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

DQ 1322 GGTGTAAGCTAATGAAGTCATCCGGCAAGTCGATGTCGTAATAACTTGCACAGGAAA 1381
|||||
|||
Db 1 GGTGTAAGCTAATGAAGTCATCCGGCAAGTCGATGTCGTAATAACTTGCACAGGAAA 60

QY 1382 TAAGATGTAGTGACAGCGGAGCACTGTCGATGAAAAACAGTTGATCGTATGCAA 1441
|
Db 61 TAAGATGTAGTGACAGCGGAGCACTGTCGATGAAAAACAGTTGATCGTATGCAA 120
QY 1442 TATGGCCACTCCACACAGAAATGATGTGACCAGCCTCCGCACTCCGAGCTGACGTG 1501
|
Db 121 TATGGCCACTCCACACAGAAATGATGTGACCAGCCTCCGCACTCCGAGCTGACGTG 180
QY 1502 GGAGCGAGTACGTCTCAGGTGGACCATGTCATCTGGCCAGATG 1545
|
Db 181 GGAGCGAGTACGTCTCAGGTGGACCATGTCATCTGGCCAGATG 224

RESULT 7

US-10-021-121-1/c
; Sequence 1, Application US/10021121
; Patent No. US2002014244A1

; GENERAL INFORMATION:

APPLICANT: Caras, Ingrid W
TITLE OF INVENTION: A2-1 Neurotrophic Factor
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 Inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/021,121
FILING DATE: 06-Dec-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/635,130
FILING DATE: 19-Mar-1996

ATTORNEY/AGENT INFORMATION:

NAME: Torchia, PhD., Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: P1001
TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/225-8674
TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1877 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Linear

FEATURE:

NAME/KEY: Extra Cellular Domain
LOCATION: 244-899

IDENTIFICATION METHOD:

OTHER INFORMATION:

FEATURE:

NAME/KEY: Transmembrane Domain
LOCATION: 901-978

IDENTIFICATION METHOD:

OTHER INFORMATION:

FEATURE:

NAME/KEY: signal peptide
LOCATION: 244-321

IDENTIFICATION METHOD:

OTHER INFORMATION:

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-10-021-121-1

Query Match

7.4%; Score 190; DB 12; Length 1877;

Best Local Similarity 100.0%; Pred. No. 2.1e-90;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1283 CCCCATCTGTGCTCGAGCCCTGCATGATGGGTTCAGGGTGTAAAGCTAAATGAAGT 1342
|
Db 1446 CCCCATCTGTGCTCGAGCCCTGCATGATGGGTTCAGGGTGTAAAGCTAAATGAAGT 1387
QY 1343 CATCCGGCAAGTCGATGTCTAATACTTGCACAGGAATAAGATGTAGTACAGCGGA 1402
|
Db 1386 CATCCGGCAAGTCGATGTCTAATACTTGCACAGGAATAAGATGTAGTACAGCGGA 1327
QY 1403 GCACTTGATCGCATGAAAAACAGTTGTATCGTATGCAATATGGCCCACTCCAACAGGA 1462
|
Db 1326 GCACTTGATCGCATGAAAAACAGTTGTATCGTATGCAATATGGCCCACTCCAACAGGA 1267
QY 1463 AATCGATGTG 1472
|
Db 1266 AATCGATGTG 1257

RESULT 8

US-09-880-107-346

; Sequence 346, Application US/09880107
; Patent No. US20020142981A1

; GENERAL INFORMATION:

APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Scherf, Uwe
APPLICANT: Gene Logic, Inc.

TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer

FILE REFERENCE: 44921-5028-WO

CURRENT APPLICATION NUMBER: US/09/880,107

PRIOR FILING DATE: 2001-06-14

PRIOR APPLICATION NUMBER: US 60/211,379

PRIOR FILING DATE: 2000-06-14

PRIOR APPLICATION NUMBER: US 60/237,054

PRIOR FILING DATE: 2000-10-02

NUMBER OF SEQ ID NOS: 3950

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 346

LENGTH: 384

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA157401

US-09-880-107-346

Query Match

6.6%; Score 168; DB 10; Length 384;

Best Local Similarity 100.0%; Pred. No. 9.5e-79;
Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1065 CTCTGTGTCCGGCCATGACGTCATGATCTGTCTACCAACAGAGTTGATACTTG 1124
|
Db 73 CTCTGTGTCCGGCCATGACGTCATGATCTGTCTACCAACAGAGTTGATACTTG 132
QY 1125 TACTGCTGCCGAGATCCATTTTGGATGGCTGAAGAGACACAGATGTGATGTTGCT 1184
|
Db 133 TACTGCTGCCGAGATCCATTTTGGATGGCTGAAGAGACACAGATGTGATGTTGCT 192
QY 1185 GGAACAAGTGTGTGTGTGCTATGTGAGGTAGGCAAGGCTGC 1232
|
Db 193 GGAACAAGTGTGTGTGTGCTATGTGAGGTAGGCAAGGCTGC 240

RESULT 9

US-09-783-590-278

; Sequence 278, Application US/09783590
; Patent No. US20020110850A1

; GENERAL INFORMATION:

APPLICANT: Dillon, Patrick J.

APPLICANT: Haseltine, William A.

APPLICANT: Li, Haodong

APPLICANT: Rosen, Craig A.

```

; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 278
; LENGTH: 140
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (71)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (117)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (125)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-783-590-278

Query Match          1.0%; Score 45; DB 10; Length 140;
Best Local Similarity 100.0%; Pred. No. 2.3e-13;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2519 CCTCTAGATTATTCTGTGTCCAAAAA 2563
DB 72 CCTCTAGATTATTCTGTGTCCAAAAA 116

RESULT 10
US-10-114-893-33
; Sequence 33, Application US/10114893
; Publication No. US20020193567A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallee, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Bowman, Michael R.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Carlin-Duckett, McKeough
; APPLICANT: Kelleher, Kerry S.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: GI 6000-10A
; CURRENT APPLICATION NUMBER: US/10/114,893
; CURRENT FILING DATE: 2002-04-02
; EARLIER APPLICATION NUMBER: 09/413,232
; EARLIER FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 33
; LENGTH: 863
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-114-893-33

Query Match          1.0%; Score 26; DB 9; Length 863;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 2538 TTGTCCAAAAA 2563
DB 797 TTGTCCAAAAA 822

RESULT 11
US-09-770-445-261/c
; Sequence 261, Application US/09770445
; Patent No. US20020023281A1
; GENERAL INFORMATION:
; APPLICANT: Goriach, Jörn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Krickler, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; TITLE OF INVENTION: thaliana
; FILE REFERENCE: 2023US (PARA-012PRV)
; CURRENT APPLICATION NUMBER: US/09/770,445
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/178,472
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 261
; LENGTH: 987
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(987)
; OTHER INFORMATION: n = A,T,C or G
; US-09-770-445-261

Query Match          1.0%; Score 26; DB 10; Length 987;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2538 TTGTCCAAAAA 2563
DB 38 TTGTCCAAAAA 13

RESULT 12
US-09-847-519A-1
; Sequence 1, Application US/09847519A
; Patent No. US20020102693A1
; GENERAL INFORMATION:
; APPLICANT: Luche, Ralf M.
; APPLICANT: Wel, Bo
; TITLE OF INVENTION: DSP-14 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.422
; CURRENT APPLICATION NUMBER: US/09/847,519A
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1165
```

TYPE: DNA
ORGANISM: Homo sapiens
US-09-847-519A-1

Query Match 1.0%; Score 26; DB 10; Length 1165;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2538 TTGTCACAAAAA 2563
Db 1139 TTGTCACAAAAA 1164

RESULT 13

US-09-938-842A-3304
Sequence 3304, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
FILE REFERENCE: SAME, AND METHODS OF USE
CURRENT APPLICATION NUMBER: US/09/938, 842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 3304
LENGTH: 2000
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-3304

Query Match 1.0%; Score 26; DB 9; Length 2000;
Best Local Similarity 100.0%; Pred. No. 0.0035;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2538 TTGTCACAAAAA 2563
Db 940 TTGTCACAAAAA 965

RESULT 14

US-09-887-576-15/C
Sequence 15, Application US/09887576
Patent No. US20020144047A1
GENERAL INFORMATION:
APPLICANT: Budworth, P.
APPLICANT: Brown, D.
APPLICANT: Chang, H.
APPLICANT: Zhu, T.
APPLICANT: Han, B.
APPLICANT: Wang, X.
APPLICANT: Cooper, Bret
TITLE OF INVENTION: Promoters for regulation of plant expression
FILE REFERENCE: 1360.001US1
CURRENT APPLICATION NUMBER: US/09/887, 576
CURRENT FILING DATE: 2001-06-25
PRIOR APPLICATION NUMBER: US 60/213,848
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: US 60/214,087
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: US 60/258,692
PRIOR FILING DATE: 2000-12-29
NUMBER OF SEQ ID NOS: 875
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 15
LENGTH: 2003
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-887-576-15

Query Match 1.0%; Score 26; DB 10; Length 2003;
Best Local Similarity 100.0%; Pred. No. 0.0035;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2538 TTGTCACAAAAA 2563
Db 1009 TTGTCACAAAAA 984

RESULT 15

US-09-783-590-4696
Sequence 4696, Application US/09783590
Patent No. US20020110850A1
GENERAL INFORMATION:
APPLICANT: Dillon, Patrick J.
APPLICANT: Haseltine, William A.
APPLICANT: Li, Haodong
APPLICANT: Rosen, Craig A.
APPLICANT: Ruben, Steven M.
TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
FILE REFERENCE: PO-16.2C1
CURRENT APPLICATION NUMBER: US/09/783, 590
CURRENT FILING DATE: 2000-02-15
PRIOR APPLICATION NUMBER: 08/420,856
PRIOR FILING DATE: 1995-04-12
PRIOR APPLICATION NUMBER: 08/346,731
PRIOR FILING DATE: 1994-11-21
NUMBER OF SEQ ID NOS: 12485
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4696
LENGTH: 101
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (10)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (89)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (91)
OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-4696

Query Match 1.0%; Score 25; DB 10; Length 101;
Best Local Similarity 100.0%; Pred. No. 0.0099;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2539 TTGTCACAAAAA 2563
Db 60 TTGTCACAAAAA 84

Search completed: April 20, 2003, 23:52:52
Job time : 292 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: Apr 11 20, 2003, 19:06:54 ; Search time 3261 Seconds
(without alignments)
12728.927 Million cell updates/sec

Title: US-09-782-051-1
Perfect score: 2563
Sequence: 1 ggcgcgcgcgcagtcgcgagct.....aaaaaaaaaaaaaaaaaaaaa 2563

Scoring table: OLIGO_NTC
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estopl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	850	33.2	938	9	AL527928 AL527928
2	837	32.7	921	9	AL551097 AL551097
3	826	32.2	1379	14	U69200 U69200 Soar
4	788	30.7	867	13	BM453920 BM453920 AGENCOURT
5	766	29.9	839	14	BQ212518 BQ212518 AGENCOURT
6	738	28.8	858	9	AL554850 AL554850

7	735	28.7	1021	14	BQ059373	BQ059373	AGENCOURT
8	719	28.1	1042	13	BM449470	BM449470	AGENCOURT
9	693	27.0	894	14	BQ896777	BQ896777	AGENCOURT
10	682	26.6	967	13	BM468564	BM468564	AGENCOURT
11	676	26.4	745	9	AU131962	AU131962	AGENCOURT
12	676	26.4	1059	13	BM547143	BM547143	AGENCOURT
13	673	26.3	891	9	AL520704	AL520704	AGENCOURT
14	671	26.2	871	12	BG708016	BG708016	AGENCOURT
15	656	25.6	933	14	BQ889173	BQ889173	AGENCOURT
16	655	25.6	1042	13	BM452485	BM452485	AGENCOURT
17	649	25.3	832	10	BE270773	BE270773	AGENCOURT
18	648	25.3	764	9	AU126038	AU126038	AGENCOURT
19	634	24.7	751	9	AU133711	AU133711	AGENCOURT
20	634	24.7	969	12	BG259448	BG259448	AGENCOURT
21	631	24.6	928	14	BQ641487	BQ641487	AGENCOURT
22	629	24.5	1098	13	BM454352	BM454352	AGENCOURT
23	626	24.4	865	9	AL519571	AL519571	AGENCOURT
24	621	24.2	748	9	AU124992	AU124992	AGENCOURT
25	621	24.2	957	9	AL520703	AL520703	AGENCOURT
26	606	23.6	700	9	AU133884	AU133884	AGENCOURT
27	606	23.6	944	12	BF033444	BF033444	AGENCOURT
28	605	23.6	673	12	BG501360	BG501360	AGENCOURT
29	600	23.4	747	12	BG697210	BG697210	AGENCOURT
30	592	23.1	598	14	BQ614546	BQ614546	AGENCOURT
31	589	23.0	589	14	BM750989	BM750989	AGENCOURT
32	589	23.0	691	14	BM837838	BM837838	AGENCOURT
33	585	22.8	927	14	BQ894873	BQ894873	AGENCOURT
34	580	22.6	1088	12	BF797186	BF797186	AGENCOURT
35	578	22.6	629	14	BM786844	BM786844	AGENCOURT
36	573	22.4	1174	14	BM927200	BM927200	AGENCOURT
37	563	22.0	590	10	BE514092	BE514092	AGENCOURT
38	563	22.0	614	14	BM837761	BM837761	AGENCOURT
39	558	21.8	744	12	BG699160	BG699160	AGENCOURT
40	556	21.7	801	13	BI551474	BI551474	AGENCOURT
41	555	21.7	683	12	BG740147	BG740147	AGENCOURT
42	554	21.6	950	9	AL527927	AL527927	AGENCOURT
43	551	21.5	667	14	BM785550	BM785550	AGENCOURT
44	547	21.3	785	12	BG699237	BG699237	AGENCOURT
45	547	21.3	846	12	BG164829	BG164829	AGENCOURT

ALIGNMENTS

RESULT 1
AL527928
LOCUS
DEFINITION AL527928 LFI_NFL003_NBC3 Homo sapiens CDNA clone CS0DC027YC19 5
ACCESSION AL527928
VERSION AL527928
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 938)
Ll.W.B., Gruber,C., Jesse,J. and Polayes,D.
Full-length CDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source
1. 938
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DC027YC19"
/clone_11b="LFI_NFL003_NBC3"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : filiang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 241 a 208 c 275 g 212 t 2 others
ORIGIN

Query Match 33.2%; Score 850; DB 9; Length 938;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 900; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 289 TCGGGGAGAGCTGAAGCAGGCGCCAGGAGATCGAGCGCCGAGAGTACTCTTCATGG 348
DB 1 TCGGGGAGAGCTGAAGCAGGCGCCAGGAGATCGAGCGCCGAGAGTACTCTTCATGG 60
OY 349 CCACCGTCACCAAGCGCCCAAGAGCAATTCAGTTGCTGATGACATGACAGAGTTCA 408
DB 61 CCACCGTCACCAAGCGCCCAAGAGCAATTCAGTTGCTGATGACATGACAGAGTTCA 120
OY 409 CCAATTCGCCCAACAAACTGGCCGAGATCTTGTCTGCTGATCTCAGTCTCACA 468
DB 121 CCAATTCGCCCAACAAACTGGCCGAGATCTTGTCTGCTGATCTCAGTCTCACA 180
OY 469 CTGACAGCTACAGTCTGATCTGATCTCAGATAGTCTGATGATGAGTTCTCCCC 528
DB 181 CTGACAGCTACAGTCTGATCTGATCTCAGATAGTCTGATGATGAGTTCTCCCC 240
OY 529 GAGAGAAGCAGCAACCACTCCAGGGGAGCAATTTCTGTGAGAGACATCAAGC 588
DB 241 GAGAGAAGCAGCAACCACTCCAGGGGAGCAATTTCTGTGAGAGACATCAAGC 300
OY 589 AGGAGAAATTTGGAGCGCCGGGAGATGAGATGAGAGAGCAAGACATGTCCTCTGATTT 648
DB 301 AGGAGAAATTTGGAGCGCCGGGAGATGAGATGAGAGAGCAAGACATGTCCTCTGATTT 360
OY 649 CACTCAGGAACGCTGCTCAGGGGAGAGCGCTGGCTGCTAAATATAGTGGCTGTA 708
DB 361 CACTCAGGAACGCTGCTCAGGGGAGAGCGCTGGCTGCTAAATATAGTGGCTGTA 420
OY 709 CACACATCACAGCCAGACAGCGGTGTGATGAGACACTCTGTGCCCTGGGGCTCAGT 768
DB 421 CACACATCACAGCCAGACAGCGGTGTGATGAGACACTCTGTGCCCTGGGGCTCAGT 480
OY 769 GCCGCTGCTGCTGTACATCTACTCACTCAGATGAGTAGCTGCAGCACTGGCTG 828
DB 481 GCCGCTGCTGCTGTACATCTACTCACTCAGATGAGTAGCTGCAGCACTGGCTG 540
OY 829 AGGCTGAGTGTGAGTGTTCGCTTGGAGGGGAGTCAAGAGTAGTCTTGTGGTGTGA 888
DB 541 AGGCTGAGTGTGAGTGTTCGCTTGGAGGGGAGTCAAGAGTAGTCTTGTGGTGTGA 600
OY 889 TTGACCGCTGTGTAACATGAGTGGTGGAGCCCAACATGATCCTGATGATGGGGAG 948
DB 601 TTGACCGCTGTGTAACATGAGTGGTGGAGCCCAACATGATCCTGATGATGGGGAG 660
OY 949 ACTTAACCACTGGGTTTATAGAAGTATCAACGTTTAAAGAGATCCAGGCATG 1008
DB 661 ACTTAACCACTGGGTTTATAGAAGTATCAACGTTTAAAGAGATCCAGGCATG 720
OY 1009 TGAAGAGAGCGTGAAGTGTGTTACAGAGCTGTATCAGCTCTCCAAGCTGGAGCTCT 1068
DB 721 TGAAGAGAGCGTGAAGTGTGTTACAGAGCTGTATCAGCTCTCCAAGCTGGAGCTCT 780
OY 1069 GTGTTCCGGCATGAAGCTCAATGATCTGTTACCAAGAGTTGATTAATCTGTACT 1128
DB 781 GTGTTCCGGCATGAAGCTCAATGATCTGTTACCAAGAGTTGATTAATCTGTACT 840

OY 1129 CTTGCCGAGATCCATTTTGATGGCCTGAAGAGACACAGATGTGATGTTGGTGGGA 1188
DB 841 GCTGCCGAGATCCATTTTGATGGCCTGAAGAGACACAGATGTGATGTTGGTGGGA 900
OY 1189 A 1189
DB 901 A 901.

RESULT 2
AL551097

LOCUS AL551097 921 bp mRNA linear EST 16-FEB-2001

DEFINITION AL551097 LTI_NFL006.PL2 Homo sapiens cDNA clone CSODI066YM07 5 prime, mRNA sequence.

ACCESSION AL551097
VERSION AL551097.1 GI:12888715

KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 921)

AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)

COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France

BP 191 91006 Evry cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers

FEATURES
Source

1. 921
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="LTI_NFL006.PL2"
/tissue_type="placenta"

/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : filiang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 217 a 231 c 302 g 170 t 1 others
ORIGIN

Query Match 32.7%; Score 837; DB 9; Length 921;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 887; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 69 GCCTGGGCTGCCGACAGCAAGCGGTGGCCACAGACCTCAGAAAGCCGAGCAGCTC 128
DB 1 GCCTGGGCTGCCGACAGCAAGCGGTGGCCACAGACCTCAGAAAGCCGAGCAGCTC 60
OY 129 GACGACAGGGCGGAGAGGGTGGCGATGCGTGTGAGAGGGCGCGCGGAGAGGC 188
DB 61 GACGACAGGGCGGAGAGGGTGGCGATGCGTGTGAGAGGGCGCGCGGAGAGGC 120
OY 189 GGGCGGGCGGAGAGGGGGAAGAGGGGGGGGGGGGCTCAGCCGCTGGCGGGCGGC 248
DB 121 GGGCGGGCGGAGAGGGGGAAGAGGGGGGGGGGGGCTCAGCCGCTGGCGGGCGGC 180
OY 249 GGGGGAATGTGATGCTGACCGCGATGCGCTGCCCGGGGCTGGGGAGAGCTGAAGCAG 308
DB 181 GGGGGAATGTGATGCTGACCGCGATGCGCTGCCCGGGGCTGGGGAGAGCTGAAGCAG 240
OY 309 GCCAAGAGATCGAGAGCGCCGAGAGTACTCTTCATGCGCCACCGTCACCAAGCGGCC 368
DB 241 GCCAAGAGATCGAGAGCGCCGAGAGTACTCTTCATGCGCCACCGTCACCAAGCGGCC 300

QY	369	AAGAAGCAAAATCCAGTTGCTGATGACATGCAGGAGTTCCACAAATTCCCCACCACAAACT	428
Db	301	AAGAAGCAAAATCCAGTTGCTGATGACATGCAGGAGTTCCACAAATTCCCCACCACAAACT	360
QY	429	GGCCGAAGATCTTTGTCTCGCTCGATCTCACAGTCCCTCCACTGCACAGTACAGTTCAGCT	488
Db	361	GGCCGAAGATCTTTGTCTCGCTCGATCTCACAGTCCCTCCACTGCACAGTACAGTTCAGCT	420
QY	489	GCATCCTACACAGATAGCTCTGATGATGAGGTTTCTCCCCAGAGAAGCAGCAAAACCAAC	548
Db	421	GCATCCTACACAGATAGCTCTGATGATGAGGTTTCTCCCCAGAGAAGCAGCAAAACCAAC	480
QY	549	TCCAAGGGCAGCAGCAATTTCTGTGTGAAGAACATCAAGAGGCGAATTTGGAGCGCCGG	608
Db	481	TCCAAGGGCAGCAGCAATTTCTGTGTGAAGAACATCAAGAGGCGAATTTGGAGCGCCGG	540
QY	609	GAGATTGAGATTGCAGAGCAAGACATGTCCTCTGATTTCACTCAGGAAACGTGCTCAG	668
Db	541	GAGATTGAGATTGCAGAGCAAGACATGTCCTCTGATTTCACTCAGGAAACGTGCTCAG	600
QY	669	GGGAGAGAGCCCTTGGCTGGTGTCTAAATAGTGGGCTGTACACATCACAGCCACAGACA	728
Db	601	GGGAGAGAGCCCTTGGCTGGTGTCTAAATAGTGGGCTGTACACATCACAGCCACAGACA	660
QY	729	GCGGTGTGATTGAGACACTCTGTGCCCTGGGGGCTCAGTCCGCTGTGCTGTGAAC	788
Db	661	GCGGTGTGATTGAGACACTCTGTGCCCTGGGGGCTCAGTCCGCTGTGCTGTGAAC	720
QY	789	ATCTACTCAACTCAGAAATGAAGTAGCTGCAGCACCTGGCTGAGGCTGGAGTTGCAGTGTTC	848
Db	721	ATCTACTCAACTCAGAAATGAAGTAGCTGCAGCACCTGGCTGAGGCTGGAGTTGCAGTGTTC	780
QY	849	GCTTGAAGGGCGAGTCAGAAGATGACTTCTGTGTGTATTGACCGCTGTGTAACATG	908
Db	781	GCTTGAAGGGCGAGTCAGAAGATGACTTCTGTGTGTATTGACCGCTGTGTAACATG	840
QY	909	GATGGGTGCAGGCCCAACATGATCCTGGATGATGGGGGAGACTTAAAC	956
Db	841	GATGGGTGCAGGCCCAACATGATCCTGGATGATGGGGGAGACTTAAAC	888

```

RESULT 3
LOCUS      U69200
DEFINITION U69200 Soares infant brain INIB Homo sapiens CDNA clone 23931, mRNA
ACCESSION  U69200
VERSION    U69200.1
KEYWORDS   sequence.
SOURCE     EST.
           human.
ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 1379)
           Volorio,S., Simon,G., Repetto,M., Cucciardi,M., Banfi,S., Borsani
           ,G., Ballabio,A. and Zollo,M.
           Sequencing analysis of forty-eight human image cDNA clones similar
           to drosophila mutant protein
           DNA Seq. 9 (5-6), 307-315 (1998)
           99452388
COMMENT    Contact: Zollo, Massimo
           Telethon Institute of Genetics and Medicine
           Via Olgettina 58, Milan, MI 20132, Italy
           Email: zollo@tigem.it.
FEATURES
Source     Location/Qualifiers
           1..1379
           /organism="Homo sapiens"
           /db_xref="taxon:9606"
           /map="1p13.3"
           /clone="23931"
           /clone_lib="Soares infant brain INIB"
           /sex="female"
           /dev_stage="73 days post natal"

```

```

/lab_host="DH10B (ampicillin resistant)"
/note="Organ: whole brain; Vector: Lactamid BA; Site_1: Not
I; Site_2: Hind III; 1st strand cDNA was primed with a Not
I - Oligo(dT) primer [5';
AACTGGAAGATTCGGCCGACAGATTTTTTTTTTTT 3'];
double-stranded cDNA was ligated to Hind III adaptors
(Pharmacia), digested with Not I and directionally cloned
into the Not I and Hind III sites of the Lactamid BA vector.
Library went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

```

Query Match	32.2%	Score 826	DB 14	Length 1379
Best Local Similarity	100.0%	Pred. No. 0		
Matches 826	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY 1201	TGTGTGCGTATGTGTGAGTGAAGCAAGGGCTGTGTGTGTCTCTCAAAAGCTCTTGAGCAA	1260		
Db 3	TGTGTGCGTATGTGTGAGTGAAGCAAGGGCTGTGTGTGTCTCTCAAAAGCTCTTGAGCAA	62		
QY 1261	TTGTCTACATTTACCGAAATCGAACCCTATCTGTGCTCTGCAAGGCTGCATGGATGGTTCA	1320		
Db 63	TTGTCTACATTTACCGAAATCGAACCCTATCTGTGCTCTGCAAGGCTGCATGGATGGTTCA	122		
QY 1321	GGGTGGTAAAGCTAAATGTAAGTTCATCCGGCAAGTCGATGTCTAATACTTGCACAGGAA	1380		
Db 123	GGGTGGTAAAGCTAAATGTAAGTTCATCCGGCAAGTCGATGTCTAATACTTGCACAGGAA	182		
QY 1381	ATAAGAATGTAGTGACACGGGAGCACCTTGGATCGCATGAAAAAACAGTTGTATCGTATGCA	1440		
Db 183	ATAAGAATGTAGTGACACGGGAGCACCTTGGATCGCATGAAAAAACAGTTGTATCGTATGCA	242		
QY 1441	ATATGGGCCACTCCAAACACAGAAATCGATGTGACACAGCCTCCGCACTCCGAGCTGACGT	1500		
Db 243	ATATGGGCCACTCCAAACACAGAAATCGATGTGACACAGCCTCCGCACTCCGAGCTGACGT	302		
QY 1501	GGGAGCGAGTACGTTCTCAGGTGAGCACCATGTCTATCTGGCCAGATGGCAAAACGAGTTGTCC	1560		
Db 303	GGGAGCGAGTACGTTCTCAGGTGAGCACCATGTCTATCTGGCCAGATGGCAAAACGAGTTGTCC	362		
QY 1561	TCCTGGCAGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTTCCACCTTGTCTCTGT	1620		
Db 363	TCCTGGCAGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTTCCACCTTGTCTCTGT	422		
QY 1621	CCATCAGACGCCACAACACAGGCTTTGGCACTGATAGACTCTATAATGCACCCGAGGGGC	1680		
Db 423	CCATCAGACGCCACAACACAGGCTTTGGCACTGATAGACTCTATAATGCACCCGAGGGGC	482		
QY 1681	GATACAAGCAGGATGTGTACTTGTCTTCTTCTTGAAGAAATGATGAATACGTTGCCAGCTTGC	1740		
Db 483	GATACAAGCAGGATGTGTACTTGTCTTCTTCTTGAAGAAATGATGAATACGTTGCCAGCTTGC	542		
QY 1741	ATCTGCCATCATTTGATGCCCACTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG	1800		
Db 543	ATCTGCCATCATTTGATGCCCACTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG	602		
QY 1801	GACTCAACAAAAATGGGCCATTCAAACCTAATTATTAAGATACTAATGACCACTACTAC	1860		
Db 603	GACTCAACAAAAATGGGCCATTCAAACCTAATTATTAAGATACTAATGACCACTACTAC	662		
QY 1861	CAAGGACCAAGTCCACCTGAACACACACACTCTAAGAAATATTTTAAAGATAACTTTTAT	1920		
Db 663	CAAGGACCAAGTCCACCTGAACACACACACTCTAAGAAATATTTTAAAGATAACTTTTAT	722		
QY 1921	TTTCTCTTACTCTCTTCTCTCTGATTTTTCCTATAAATTTTCATCTCTGTTTTCATC	1980		
Db 723	TTTCTCTTACTCTCTTCTCTCTGATTTTTCCTATAAATTTTCATCTCTGTTTTCATC	782		
QY 1981	TCATTAATCCAAGTTCTGACAGACACACAGGAAGTCTGCTCATGGCT 2026			
Db 783	TCATTAATCCAAGTTCTGACAGACACACAGGAAGTCTGCTCATGGCT 828			

```

RESULT 4
BM453920
LOCUS      BM453920
DEFINITION AGENCOURT_6402958 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5528995
5', mRNA sequence.
ACCESSION  BM453920
VERSION     BM453920.1
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 867)
AUTHORS   NIH-MGC http://mhc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgabbs-remail.nih.gov
          Tissue Procurement: ATCC/DCTD/DTP
          CDNA Library Preparation: Life Technologies, Inc.
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM12206 row: 1 column: 20
          High quality sequence stop: 670.
FEATURES
source
1..867
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5528995"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: PCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT  222 a 201 c 241 g 202 t 1 others
ORIGIN
Query Match 30.7%; Score 788; DB 13; Length 867;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 788; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 328 CCGAGAAGTACTCTTCATGCGCCACCGCTCACCAAGCGGCCCAAGAAGCAATCCAGTTTG 387
DB 1 CCGAGAAGTACTCTTCATGCGCCACCGCTCACCAAGCGGCCCAAGAAGCAATCCAGTTTG 60
QY 388 CTGATGACATGCAGAGGTTTACCAAAATTCCTCCACCAAAACTGGCCGAAGATCTTGTCTC 447
DB 61 CTGATGACATGCAGAGGTTTACCAAAATTCCTCCACCAAAACTGGCCGAAGATCTTGTCTC 120
QY 448 GCTCGATCTCACAGTCTCTCACTGACAGCTACAGTTGAGCTGATCTTACACAGATAGCT 507
DB 121 GCTCGATCTCACAGTCTCTCACTGACAGCTACAGTTGAGCTGATCTTACACAGATAGCT 180
QY 508 CTGATGATGAGGTTCTCTCCCGAGAGAGAGCAAACTCCAAAGGGCAGACGAATT 567
DB 181 CTGATGATGAGGTTCTCTCCCGAGAGAGAGCAAACTCCAAAGGGCAGACGAATT 240
QY 568 TCTGTGTGAAGACATCAAGCAGGAGATTTGACGCCGGGAGATTGAGATTGACAGC 627
DB 241 TCTGTGTGAAGACATCAAGCAGGAGATTTGACGCCGGGAGATTGAGATTGACAGC 300
QY 628 AAGACATGTCGTCCTGATTTCACTCAGGAACGTCGTCAGGGGGAGAGAGCCCTTGCTG 687
DB 301 AAGACATGTCGTCCTGATTTCACTCAGGAACGTCGTCAGGGGGAGAGAGCCCTTGCTG 360
QY 688 GTGCTAAATAGTGGCTGTACACACATCAGAGCCGAGAGAGCGGTTGATTGAGACAC 747

```

```

DB 361 GTGCTAAATAGTGGCTGTACACACATCAGAGCCGAGAGAGCGGTTGATTGAGACAC 420
QY 748 TCTGTGCCCTGGGGGCTCAGTGCCGCTGGTCTGCTGTACATCTACTCAACTCAGATG 807
DB 421 TCTGTGCCCTGGGGGCTCAGTGCCGCTGGTCTGCTGTACATCTACTCAACTCAGATG 480
QY 808 AAGTAGCTGACAGACTGCTGAGGCTGAGTTGACGTTGCTTGAAGGGCGAGTCAG 867
DB 481 AAGTAGCTGACAGACTGCTGAGGCTGAGTTGACGTTGCTTGAAGGGCGAGTCAG 540
QY 868 AAGTAGCTTCTGTTGTATTTGACCCGCTGTGTGAACATGATGGGTGCGAGGCCAACA 927
DB 541 AAGTAGCTTCTGTTGTATTTGACCCGCTGTGTGAACATGATGGGTGCGAGGCCAACA 600
QY 928 TGATCCTGATGATGGGGGAGACTTAACCCACTGGGTTTATAAGAGTATCCAAACGTG 987
DB 601 TGATCCTGATGATGGGGGAGACTTAACCCACTGGGTTTATAAGAGTATCCAAACGTG 660
QY 988 TTAAGAAGATCCGAGGATTTGGAAGAGAGCGTGACTGGTGTTCACAGGCTGTATCAGC 1047
DB 661 TTAAGAAGATCCGAGGATTTGGAAGAGAGCGTGACTGGTGTTCACAGGCTGTATCAGC 720
QY 1048 TCTCCAAGCTGGGAGCTCTGTCTTCCGGCCATGAACGTCATGATTTCTGTACCAAC 1107
DB 721 TCTCCAAGCTGGGAGCTCTGTCTTCCGGCCATGAACGTCATGATTTCTGTACCAAC 780
QY 1108 AGAAGTTT 1115
DB 781 AGAAGTTT 788
RESULT 5
BQ212518
LOCUS      BQ212518
DEFINITION AGENCOURT_7595081 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6062515
5', mRNA sequence.
ACCESSION  BQ212518
VERSION     BQ212518.1
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 839)
AUTHORS   NIH-MGC http://mhc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgabbs-remail.nih.gov
          Tissue Procurement: ATCC/DCTD/DTP
          CDNA Library Preparation: Life Technologies, Inc.
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM13334 row: b column: 20
          High quality sequence stop: 671.
FEATURES
source
1..839
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6062515"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: PCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT  228 a 203 c 178 g 229 t 1 others
ORIGIN

```



```
|||||
DB 726 CTGGTTTATAGAGATATCCAAACGTGTTTAGAGAGATCCGAGCATGTGAGAGAGAG 785
OY 1019 CGTACTGCTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGAGCTGTGTTCGGC 1078
|||||
DB 786 CGTACTGCTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGAGCTGTGTTCGGC 845
OY 1079 CATGAACGT 1087
|||||
DB 846 CATGAACGT 854

RESULT 7
BO059373
LOCUS BO059373 1021 bp mRNA linear EST 29-MAR-2002
DEFINITION AGENCOURT_7050082 NIH_MGC_99 Homo sapiens CDNA clone IMAGE:5815413
5', mRNA sequence.
ACCESSION BO059373
VERSION BO059373.1 GI:19818713
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1021)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Lou Staudt
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM2070 row: j column: 22
High quality sequence stop: 688.
Location/Qualifiers
1.1021
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5815413"
/clone_lib="NIH_MGC_99"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-CDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

BASE COUNT 284 a 245 c 201 g 291 t
ORIGIN
Query Match 28.7%; Score 735; DB 14; Length 1021;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1260 ATTGCTACATTACCGAATCGACCCCATCTGCTCTGCAGGCTGCATGATGGGTTTC 1319
|||||
DB 1 ATTGCTACATTACCGAATCGACCCCATCTGCTCTGCAGGCTGCATGATGGGTTTC 60
OY 1320 AGGTGTAAGCTAAATGAAGTCATCCGCGAAGTCGATGCTAATACTGCACAGGA 1379
|||||
DB 61 AGGTGTAAGCTAAATGAAGTCATCCGCGAAGTCGATGCTAATACTGCACAGGA 120
OY 1380 AATAAGATGTAGTACACGAGGACCTTGATCGCATGAAAACAGTGTATCGTATGC 1439
|||||
```

```
DB 121 AATAAGATGTAGTACACGAGGACCACTTGATCGCATGAAAACAGTGTATCGTATGC 180
OY 1440 AATATGGCCACTCCAAACAGAAATCGATGTGACCAAGCTCCGCACTCCGAGCTGACG 1499
|||||
DB 181 AATATGGCCACTCCAAACAGAAATCGATGTGACCAAGCTCCGCACTCCGAGCTGACG 240
OY 1500 TGGAGCGAGTACGTTCTCAGGTGCAACATGTATCTGGCCAGATGGCAACAGGTTGTC 1559
|||||
DB 241 TGGAGCGAGTACGTTCTCAGGTGCAACATGTATCTGGCCAGATGGCAACAGGTTGTC 300
OY 1560 CTCCTGGCAGAGGCTCGTCTACTCAATTGAGCTGCTCCACAGTTCCACCTTGTCTG 1619
|||||
DB 301 CTCCTGGCAGAGGCTCGTCTACTCAATTGAGCTGCTCCACAGTTCCACCTTGTCTG 360
OY 1620 TCATACAGCCACCAACAGGCTTTGGCACTGATAGAACTCTATATGCAACCCGAGGG 1679
|||||
DB 361 TCATACAGCCACCAACAGGCTTTGGCACTGATAGAACTCTATATGCAACCCGAGGG 420
OY 1680 CGATACAGCAGGATGTGTACTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1739
|||||
DB 421 CGATACAGCAGGATGTGTACTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 480
OY 1740 CATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTG 1799
|||||
DB 481 CATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTG 540
OY 1800 GCACTCAACAAAATGGGCCATTTCAACCTAATTATTACAGATACTAATGACCATACTA 1859
|||||
DB 541 GCACTCAACAAAATGGGCCATTTCAACCTAATTATTACAGATACTAATGACCATACTA 600
OY 1860 CCAAGGACCACTCCACCTGACACACACACTCTAAGAAATATTTTAAAGATAACTTTTA 1919
|||||
DB 601 CCAAGGACCACTCCACCTGACACACACACTCTAAGAAATATTTTAAAGATAACTTTTA 660
OY 1920 TTTCTCTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1979
|||||
DB 661 TTTCTCTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
OY 1980 CTCATTATCCAAGTT 1994
|||||
DB 721 CTCATTATCCAAGTT 735

RESULT 8
BM449470
LOCUS BM449470 1042 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6400881 NIH_MGC_67 Homo sapiens CDNA clone IMAGE:5493677
5', mRNA sequence.
ACCESSION BM449470
VERSION BM449470.1 GI:18498510
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1042)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM12117 row: m column: 06
High quality sequence stop: 694.
Location/Qualifiers
1.1042
/organism="Homo sapiens"
```

```

/db_xref="taxon:9606"
/clone="IMAGE:5493677"
/clone_1lb="NIH_MGC_67"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: PCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."
BASE COUNT      267 a      219 c      309 g      240 t      7 others
ORIGIN

Query Match      28.1%; Score 719; DB 13; Length 1042;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 769; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 532 AGAAGCAGCAACCACTCCAGGGCAGACGAATTTCTGTGTAAGAACAATCAAGCAGG 591
    |||||||
DB 1 AGAAGCAGCAACCACTCCAGGGCAGACGAATTTCTGTGTAAGAACAATCAAGCAGG 60

QY 592 CAGAATTGGACGCCGGGAGATTGAGATTGCAGACAGACATGCTGCTGTGATTTCAC 651
    |||||||
DB 61 CAGAATTGGACGCCGGGAGATTGAGATTGCAGACAGACATGCTGCTGTGATTTCAC 120

QY 652 TCAGGAACGCTGCTCAGGGGAGAGAGCCCTTGCTGCTGCTAAATAGTGGGCTGTACAC 711
    |||||||
DB 121 TCAGGAACGCTGCTCAGGGGAGAGAGCCCTTGCTGCTGCTAAATAGTGGGCTGTACAC 180

QY 712 ACATCACAGCCCAAGACAGCGGTGTTGATTGAGACACTCTGTGCCCTGGGGCTCAGTCC 771
    |||||||
DB 181 ACATCACAGCCCAAGACAGCGGTGTTGATTGAGACACTCTGTGCCCTGGGGCTCAGTCC 240

QY 772 GCTGCTGCTGTGATCACTCACTCAGAAAGTAGCTGCAGCACTGGCTGAGG 831
    |||||||
DB 241 GCTGCTGCTGTGATCACTCACTCAGAAAGTAGCTGCAGCACTGGCTGAGG 300

QY 832 CTGAGTGTGACGTGCTGCTGGAAGGGCAGTCAAGATGACTTCTGTGTGTATG 891
    |||||||
DB 301 CTGAGTGTGACGTGCTGCTGGAAGGGCAGTCAAGATGACTTCTGTGTGTATG 360

QY 892 ACCGCTGTGTAACATGATGGGTGGAGCCCAACATGATCCTGATGGGGAGACT 951
    |||||||
DB 361 ACCGCTGTGTAACATGATGGGTGGAGCCCAACATGATCCTGATGGGGAGACT 420

QY 952 TAACCCACTGGGTTTAAGAAGTATCCAAACGTGTTTAAGAAGATCCGAGGCATGTGG 1011
    |||||||
DB 421 TAACCCACTGGGTTTAAGAAGTATCCAAACGTGTTTAAGAAGATCCGAGGCATGTGG 480

QY 1012 AAGAGCGGTGACGTGCTTACAGGCTGATCAGCTCTCCAAAGCTGGAGCTGTGG 1071
    |||||||
DB 481 AAGAGCGGTGACGTGCTTACAGGCTGATCAGCTCTCCAAAGCTGGAGCTGTGG 540

QY 1072 TTCGGGCATGAACGTCATGATTTCTGTACCAACAGAGTTTGATACTGTACTGCT 1131
    |||||||
DB 541 TTCGGGCATGAACGTCATGATTTCTGTACCAACAGAGTTTGATACTGTACTGCT 600

QY 1132 GCCGAGATCCATTTGGATGGCCTGAAGAGGACACAGATGTGATTTGGTGGGAAC 1191
    |||||||
DB 601 GCCGAGATCCATTTGGATGGCCTGAAGAGGACACAGATGTGATTTGGTGGGAAC 660

QY 1192 AAGTGGTGTGTGCTGATGTGAGTAGGCAAGGCTGCTGCTCTCAAGCTC 1251
    |||||||
DB 661 AAGTGGTGTGTGCTGATGTGAGTAGGCAAGGCTGCTGCTCTCAAGCTC 720

QY 1252 TTGGAGCAATGTCTACATTAACGAATGACCCCATCTGTCTCTGCAG 1301
    |||||||
DB 721 TTGGAGCAATGTCTACATTAACGAATGACCCCATCTGTCTCTGCAG 770

RESULT 9
BO896777      894 bp      mRNA      linear      EST 16-AUG-2002
LOCUS      BO896777
DEFINITION      AGENCOURT_8715624 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6425821
```

```

5' mRNA sequence.
ACCESSION      BO896777
VERSION      BO896777.1 GI:22288791
KEYWORDS
SOURCE
ORGANISM      Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 894)
                NIH-MGC http://mgc.nci.nih.gov/.
                National Institutes of Health, Mammalian Gene Collection (MGC)
                Unpublished (1999)
                Contact: Robert Strausberg, Ph.D.
                Email: cgabs-remail.nih.gov
                Tissue Procurement: ATCC
                CDNA Library Preparation: Rubin Laboratory
                CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
                DNA Sequencing by: Agencourt Bioscience Corporation
                Clone distribution: MGC clone distribution information can be
                found through the I.M.A.G.E. Consortium/LNL at:
                http://image.lnl.gov
                Plate: LCM2608 row: d column: 14
                High quality sequence stop: 700.
                Location/Qualifiers
                source
                1. 894
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:6425821"
                /clone_1lb="NIH_MGC_101"
                /tissue_type="epidermoid carcinoma, cell line"
                /lab_host="DH10B (phage-resistant)"
                /note="Organ: lung; Vector: pOTB7; Site_1: EcoRI; Site_2:
                XhoI; cDNA made by oligo-dT priming. Directionally cloned
                into EcoRI/XhoI sites using the following 5' adaptor:
                GGACAGAG(G). Library constructed by Ling Hong in the
                laboratory of Gerald M. Rubin (University of California,
                Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                Superscript II RT (Life Technologies). Note: this is a
                NIH_MGC Library."
BASE COUNT      226 a      205 c      247 g      216 t
ORIGIN

Query Match      27.0%; Score 693; DB 14; Length 894;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 693; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 910 ATGGGTGCGAGGCCAACATGATCTGATGATGGGGAGACTTAACCACTGGTTATA 969
    |||||||
DB 1 ATGGGTGCGAGGCCAACATGATCTGATGATGGGGAGACTTAACCACTGGTTATA 60

QY 970 AGAATATCCAAACGTGTTTAAGAAGATCCGAGCATTTGTGAAGAGCGTGAAGTG 1029
    |||||||
DB 61 AGAATATCCAAACGTGTTTAAGAAGATCCGAGCATTTGTGAAGAGCGTGAAGTG 120

QY 1030 TTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTGTTCGGCCATGAAGTCA 1089
    |||||||
DB 121 TTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTGTTCGGCCATGAAGTCA 180

QY 1090 ATGATTCGTGTACCAACAGAGTTGATTAATCTGTACTGCTGCCAGAAATCCATTGG 1149
    |||||||
DB 181 ATGATTCGTGTACCAACAGAGTTGATTAATCTGTACTGCTGCCAGAAATCCATTGG 240

QY 1150 ATGGCTGAAGAGGACACAGATGTGATGTTGGTGGGAACAAGTGGTGTGGCT 1209
    |||||||
DB 241 ATGGCTGAAGAGGACACAGATGTGATGTTGGTGGGAACAAGTGGTGTGGCT 300

QY 1210 ATGTTAGGTAGGCAAGGGCTGCTGCTCTCAAGCTCTTGAGCAATGTGCTACA 1269
    |||||||
DB 301 ATGTTAGGTAGGCAAGGGCTGCTGCTCTCAAGCTCTTGAGCAATGTGCTACA 360

QY 1270 TTACCGAATGACCCCATCTGTGCTGTGCGAGCCCTGCATGATGGGTGAGGTGTA 1329
    |||||||
DB 361 TTACCGAATGACCCCATCTGTGCTGTGCGAGCCCTGCATGATGGGTGAGGTGTA 420
```

QY 1330 AGCTAATGAAGTCATCCGGCAAGTCGATGTCGTAATACTTGACAGGAATAAGATG 1389
|||||
DB 421 AGCTAATGAAGTCATCCGGCAAGTCGATGTCGTAATACTTGACAGGAATAAGATG 480
QY 1390 TAGTGACACGGGAGCAGCTTGATGCGCATGAAAAACAGTTGTATCGTATGCAATATGGGCC 1449
|||||
DB 481 TAGTGACACGGGAGCAGCTTGATGCGCATGAAAAACAGTTGTATCGTATGCAATATGGGCC 540
QY 1450 ACTCCACACAGAAATCGATGTGACACAGCCTCCGCACTCCGAGCTGACGTGGAGCGAG 1509
|||||
DB 541 ACTCCACACAGAAATCGATGTGACACAGCCTCCGCACTCCGAGCTGACGTGGAGCGAG 600
QY 1510 TACGTTCTCAGGTGGACCATGTCTATCTGGCCAGATGGCAACAGATTGTCTCTCTGGCAG 1569
|||||
DB 601 TACGTTCTCAGGTGGACCATGTCTATCTGGCCAGATGGCAACAGATTGTCTCTCTGGCAG 660
QY 1570 AGGTCGTCCTACTCAATTGAGCTGCTCCACAG 1602
|||||
DB 661 AGGTCGTCCTACTCAATTGAGCTGCTCCACAG 693

RESULT 10
LOCUS BM468564 967 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6475681 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5578170
5', mRNA sequence.
ACCESSION BM468564
VERSION BM468564.1 GI:18517606
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 967)
NIH-MGC http://mgs.ncl.nih.gov/.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DRP
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM12332 row: m column: 19
High quality sequence stop: 685.
FEATURES
source location/Qualifiers
1..967
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5578170"
/clone_1ib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: PCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."

BASE COUNT 214 a 253 c 328 g 172 t
ORIGIN
Query Match 26.6%; Score 682; DB 13; Length 967;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 782; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGCGGGGAGCTCGAGCTCGAGCTGCTGCTTGTGTTCTTGTGGCCACCCTGCGT 60
|||||
DB 3 GGGCGGGGAGCTCGAGCTCGAGCTGCTGCTTGTGTTCTTGTGGCCACCCTGCGT 62
QY 61 GTCCGGCTGCTTGGGCTGCCGACAGCAAGCGGTGGCCACAGCACTCAGAAAGCCGA 120

DB 63 GTCCGGCTGCTTGGGCTGCCGACAGCAAGCGGTGGCCACAGCACTCAGAAAGCCGA 122
QY 121 CGCAGCTCGACGAGGGGGCCGGCAGAGGGGTGGCGATCGCGTGTGGAGGGGGCCGCC 180
DB 123 CGCAGCTCGACGAGGGGGCCGGCAGAGGGGTGGCGATCGCGTGTGGAGGGGGCCGCC 182
QY 181 GGGCAGGGGGGGGGCCAGAGGGGGAAGAGGGGGGGGGGGGGGTACGGCTGGGCC 240
DB 183 GGGCAGGGGGGGGGCCAGAGGGGGAAGAGGGGGGGGGGGGGGTACGGCTGGGCC 242
QY 241 GGGCCGGGGGGGAATGTCGATGCTGACGCGATGCGCGCTGCCGGGGTGGGGAGAGC 300
DB 243 GGGCCGGGGGGGAATGTCGATGCTGACGCGATGCGCGCTGCCGGGGTGGGGAGAGC 302
QY 301 TGAAGCAGGGCCAGAGATCGAGAGCGCCGAGAGTACTCTCTCATGCGCCACCTCACCA 360
DB 303 TGAAGCAGGGCCAGAGATCGAGAGCGCCGAGAGTACTCTCTCATGCGCCACCTCACCA 362
QY 361 AGCGGCCCAAGAGCAATCCAGTTGCTGATGACATGCAAGAGTTCACCAATTCGCCA 420
DB 363 AGCGGCCCAAGAGCAATCCAGTTGCTGATGACATGCAAGAGTTCACCAATTCGCCA 422
QY 421 CCAAACTGGGGGAGAGATCTTGTCTGCTGCTGATCAGATCCTCCACTGACAGCTACA 480
DB 423 CCAAACTGGGGGAGAGATCTTGTCTGCTGCTGATCAGATCCTCCACTGACAGCTACA 482
QY 481 GTTCAGCTGATCTTACACAGATAGCTCTGATGATGAGGTTTCTCCCGAGAGAGCAGC 540
DB 483 GTTCAGCTGATCTTACACAGATAGCTCTGATGATGAGGTTTCTCCCGAGAGAGCAGC 542
QY 541 AAACCACTCCAAGGGGAGAGCAATTTCTGTGTAAGAACATCAAGCAGAGCAATTTG 600
DB 543 AAACCACTCCAAGGGGAGAGCAATTTCTGTGTAAGAACATCAAGCAGAGCAATTTG 602
QY 601 GACGCCGGAGATGATGATGTCAGAGCAGACATGCTGCTGATTTCACTCAGAGAAC 660
DB 603 GACGCCGGAGATGATGATGTCAGAGCAGACATGCTGCTGATTTCACTCAGAGAAC 662
QY 661 GTGCTCAGGGGAGAGAGCCCTTGCTGCTGCTAAATAGTGGGCTGTACACACATCAG 720
DB 663 GTGCTCAGGGGAGAGAGCCCTTGCTGCTGCTAAATAGTGGGCTGTACACACATCAG 722
QY 721 CCCAGACAGCGGTGTTGATGAGACACTGTGCGCTGGGGGCTCAGTGCGGCTGCTG 780
DB 723 CCCAGACAGCGGTGTTGATGAGACACTGTGCGCTGGGGGCTCAGTGCGGCTGCTG 782
QY 781 CTTG 784
DB 783 CTTG 786

RESULT 11
LOCUS AUI31962 745 bp mRNA linear EST 01-AUG-2002
DEFINITION AUI31962 NT2RP3 Homo sapiens cDNA clone NT2RP3003545 5', mRNA
sequence.
ACCESSION AUI31962
VERSION AUI31962.1 GI:10992316
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 745)
Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,T.
TITLES HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao. Isogai
Genomics Laboratory
Helix Research Institute

1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986

Email: genomics@hri.co.jp
HRI human cDNA project; 5' - & 3' - end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.

FEATURES

source
1. 745
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RP3003545"
/clone_lib="NT2RP3"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
cells after 2-weeks retinoic acid (RA) induction"

BASE COUNT 193 a 149 c 214 g 186 t 3 others

ORIGIN

Query Match 26.4%; Score 676; DB 9; Length 745;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 676; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 777 TCTGCTGTACATCTACTCACTCAGAATGAGTAGCTGCAGCAGCTGGCTGAGCTGGA 836
DB 1 TCTGCTGTACATCTACTCACTCAGAATGAGTAGCTGCAGCAGCTGGCTGAGCTGGA 60
OY 837 GTTGCACTGTGCTTGAAGGGCGAGTCAGAGATGACTTCTGCTGCTGATGACCGC 896
DB 61 GTTGCACTGTGCTTGAAGGGCGAGTCAGAGATGACTTCTGCTGCTGATGACCGC 120
OY 897 TGTGTGAACATGAGTGGTGGCAGGCAACATGATCTGATGATGGGAGACTTAAC 956
DB 121 TGTGTGAACATGAGTGGTGGCAGGCAACATGATCTGATGATGGGAGACTTAAC 180
OY 957 CACTGGGTATTAAGAAGTATCCAAACGTGTTAAGAAGATCCGAGCATGTGGAAG 1016
DB 181 CACTGGGTATTAAGAAGTATCCAAACGTGTTAAGAAGATCCGAGCATGTGGAAG 240
OY 1017 AGCGTACTGCTGTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTTCCG 1076
DB 241 AGCGTACTGCTGTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTTCCG 300
OY 1077 GCCATGAACGTCATGATCTGTATCCAAACAGATTGATTAAGTCTGCTGCGCA 1136
DB 301 GCCATGAACGTCATGATCTGTATCCAAACAGATTGATTAAGTCTGCTGCGCA 360
OY 1137 GAATCCATTTGATGCGCTGAAGAGACACAGATGTGATGTTGGTGGAAACAAGTG 1196
DB 361 GAATCCATTTGATGCGCTGAAGAGACACAGATGTGATGTTGGTGGAAACAAGTG 420
OY 1197 GTGCTGTGCTATGCTGAGGTAGGCAAGGCTGCTGCTGCTCTCAAGCTCTTGA 1256
DB 421 GTGCTGTGCTATGCTGAGGTAGGCAAGGCTGCTGCTGCTCTCAAGCTCTTGA 480
OY 1257 GCAATGTCTACATTAACCGAAATCGACCCCATCTGCTCTGCAGAGGCTGCATGATGG 1316
DB 481 GCAATGTCTACATTAACCGAAATCGACCCCATCTGCTCTGCAGAGGCTGCATGATGG 540
OY 1317 TTCTGGGTGTAAGCTAAATGAAGTCATCCGCAAGTGCATGCTAATACTTGACA 1376
DB 541 TTCTGGGTGTAAGCTAAATGAAGTCATCCGCAAGTGCATGCTAATACTTGACA 600
OY 1377 GGAATAAAGATAGTAGACACGGAGCAGCTTGATGCGATGAAACAGTTGATCGTA 1436
DB 601 GGAATAAAGATAGTAGACACGGAGCAGCTTGATGCGATGAAACAGTTGATCGTA 660
OY 1437 TGCATATATGGCCACT 1452
DB 661 TGCATATATGGCCACT 676

RESULT 12
BM547143 1059 bp mRNA linear EST 20-FEB-2002
LOCUS
DEFINITION
AGENCOURT_6499536 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5730185
5', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
EST.
BM547143.1 GI:18780677

ORGANISM

human.

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
NIH-MGC <http://mgc.ncl.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Invitrogen

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: L1AM12728 row: k column: 18
High quality sequence stop: 747.

FEATURES

source

Location/Qualifiers

1. 1059
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5730185"
/clone_lib="NIH_MGC_124"
/tissue_type="hippocampus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: EcoRV
(destroyed); Site_2: NotI; RNA source male hippocampus,
age 27. Library is oligo-dT primed and directionally
cloned (EcoRV site is destroyed upon cloning). Average
insert size 1.4 kb, insert size range 0.9-4 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 012."

BASE COUNT 290 a 242 c 221 g 304 t 2 others

ORIGIN

Query Match 26.4%; Score 676; DB 13; Length 1059;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 676; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1377 GGAATAAAGATAGTAGACACGGAGCAGCTTGATCGCATGAAACAGTTGATCGTA 1436
DB 20 GGAATAAAGATAGTAGACACGGAGCAGCTTGATCGCATGAAACAGTTGATCGTA 79
OY 1437 TGCATATATGGCCACTCCACACACAGAAATCGATGTGACCAAGCTCCGAGCTG 1496
DB 80 TGCATATATGGCCACTCCACACACAGAAATCGATGTGACCAAGCTCCGAGCTG 139
OY 1497 ACGTGGAGCGAGTACGTTCTCAGGTGACCATGTCACTGGCCAGATGGCAACGAGTT 1556
DB 140 ACGTGGAGCGAGTACGTTCTCAGGTGACCATGTCACTGGCCAGATGGCAACGAGTT 199
OY 1557 GTCTCTCCGAGAGGCTGCTCAATTTGAGCTGCTCCACAGTTCCACCTTTGTT 1616
DB 200 GTCTCTCCGAGAGGCTGCTCAATTTGAGCTGCTCCACAGTTCCACCTTTGTT 259
OY 1617 CTGTCCATCACAGCCACACACAGGCTTGGCACTGATAGAACTCTAATGACCCGAG 1676
DB 260 CTGTCCATCACAGCCACACACAGGCTTGGCACTGATAGAACTCTAATGACCCGAG 319
OY 1677 GGGCATACAGCAGGATGTGTACTTCTCTAAGAAATGATGAATACGTGCGAGC 1736
DB 1677 GGGCATACAGCAGGATGTGTACTTCTCTAAGAAATGATGAATACGTGCGAGC 1736

Db 320 GGGCGATACAGCAGGATGTGTACTTCTTCTTAAGAAATGGATGATACGTTGCCAGC 379
 QY 1737 TTGCATCTGCCATCATTTGATGCCCCCCTTACAGAGCTGACAGATGACCAACCAATAT 1796
 Db 380 TTGCATCTGCCATCATTTGATGCCCCCCTTACAGAGCTGACAGATGACCAACCAATAT 439
 QY 1797 CTGGACTCAACAAAATGGGCAATCAAACTTAATATTATACAGATTAATGACCATA 1856
 Db 440 CTGGACTCAACAAAATGGGCAATCAAACTTAATATTATACAGATTAATGACCATA 499
 QY 1857 CTACCAAGGAGCAGTCCACCTGAACACACACTCTAAGAATATTTTAAAGATACT 1916
 Db 500 CTACCAAGGAGCAGTCCACCTGAACACACACTCTAAGAATATTTTAAAGATACT 559
 QY 1917 TTATTTCTTCTTACTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1976
 Db 560 TTATTTCTTCTTACTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 619
 QY 1977 CATCTCATTAATCCAAAGTCTGACAGACACAGGAAGTCTGATGAGCTCTTATGATGA 2036
 Db 620 CATCTCATTAATCCAAAGTCTGACAGACACAGGAAGTCTGATGAGCTCTTATGATGA 679
 QY 2037 AATAGAGTTCAGGGT 2052
 Db 680 AATAGAGTTCAGGGT 695

RESULT 13

AL520704 891 bp mRNA linear EST 13-FEB-2001
 LOCUS AL520704 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB002YL22 5
 DEFINITION prime, mRNA sequence.

ACCESSION AL520704
 VERSION AL520704.1 GI:12784197
 KEYWORDS EST.
 SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 891)

L.I.W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

COMMENT

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

SOURCE

1.891

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS0DB002YL22"

/clone_1lb="LTI_NFL004_NBC2"

/sex="male"

/tissue_type="neuroblastoma cells"

/lab_host="DH10B"

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed

by Life Technologies. Contact : Feng Liang Life

Technologies, a division of Invitrogen 9800 Medical Center

Drive Rockville, Maryland 20850, USA Fax : (1) 301 610

8371 Email : filiang@lifetech.com URL :

http://fulllength.invitrogen.com"

BASE COUNT 194 a 238 c 299 g 158 t 2 others

ORIGIN

Query Match 26.3%; Score 673; DB 9; Length 891;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 873; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGGCGGGCAGGTCGAGACCTCGAGCTGCTGCTTCTGTTCTTGTGGCCACCCTGCT 60
 Db 15 GGGCGGGCAGGTCGAGACCTCGAGCTGCTGCTTCTGTTCTTGTGGCCACCCTGCT 74
 QY 61 GTCCGGCTGCTTGGGCTGCCGAGACAGACAGAGGCGTGGCCACAGCACCCTCAGAGCCGA 120
 Db 75 GTCCGGCTGCTTGGGCTGCCGAGACAGACAGAGGCGTGGCCACAGCACCCTCAGAGCCGA 134
 QY 121 CGCAGCTCGACGCGAGGCGCCGCGAGAGGCTGGCCATCGCGTGTGAGAGGCGCGCGCC 180
 Db 135 CGCAGCTCGACGCGAGGCGCCGCGAGAGGCTGGCCATCGCGTGTGAGAGGCGCGCGCC 194
 QY 181 GGGCAGGCGGGCGGGCGCGAGAGGCGGGGAGAGGCGGGCGGGGTGAGCCGCTGGCC 240
 Db 195 GGGCAGGCGGGCGGGCGCGAGAGGCGGGGAGAGGCGGGCGGGGTGAGCCGCTGGCC 254
 QY 241 GGGCGGGCGGGGAGATGTCGATGCTGACGCGGATGCGGCTGCGGGGTGCGGAGAGC 300
 Db 255 GGGCGGGCGGGGAGATGTCGATGCTGACGCGGATGCGGCTGCGGGGTGCGGAGAGC 314
 QY 301 TGAAGCAGGCGCAAGAGATCGAGAGCGCGGAGAGTACTCTTCATGCGCCACCGTCACCA 360
 Db 315 TGAAGCAGGCGCAAGAGATCGAGAGCGCGGAGAGTACTCTTCATGCGCCACCGTCACCA 374
 QY 361 AGCGCGCCAGAGCAATCCAGTTTCTGATGACATGACAGAGTTCACCAATTCGCCA 420
 Db 375 AGCGCGCCAGAGCAATCCAGTTTCTGATGACATGACAGAGTTCACCAATTCGCCA 434
 QY 421 CCAAACTGGCCGAGATCTTGTCTCGCTCGATCTCAGATCTCCACTGACAGCTACA 480
 Db 435 CCAAACTGGCCGAGATCTTGTCTCGCTCGATCTCAGATCTCCACTGACAGCTACA 494
 QY 481 GTTCAGCTGATCTTACACAGATAGCTCTGATGATGAGGTTTCTCCCGAGAGAGCAGC 540
 Db 495 GTTCAGCTGATCTTACACAGATAGCTCTGATGATGAGGTTTCTCCCGAGAGAGCAGC 554
 QY 541 AAACCACTCCAAAGGCGAGCAATTTCTGTGAGAAACATCAAGCAGGAGATTG 600
 Db 555 AAACCACTCCAAAGGCGAGCAATTTCTGTGAGAAACATCAAGCAGGAGATTG 614
 QY 601 GAGCGCGGAGATGAGATTGACAGAGCAAGATGCTGCTGTGATTTCACTCAGGAAC 660
 Db 615 GAGCGCGGAGATGAGATTGACAGAGCAAGATGCTGCTGTGATTTCACTCAGGAAC 674
 QY 661 GTGCTAGGCGGAGAGAGCCCTTGCTGTGCTTAATAAGTGGGCTGACACACATCAGAG 720
 Db 675 GTGCTAGGCGGAGAGAGCCCTTGCTGTGCTTAATAAGTGGGCTGACACACATCAGAG 734
 QY 721 CCCAGACAGCGGTGTTGATTGAGACACTCTGTGCTGGGGCTCAGTGCCGCTGCTG 780
 Db 735 CCCAGACAGCGGTGTTGATTGAGACACTCTGTGCTGGGGCTCAGTGCCGCTGCTG 794
 QY 781 CTTGTAACATCTACTCAACTCAGATGAGTACTGACAGCACTGGCTGAGGCTGAGTTG 840
 Db 795 CTTGTAACATCTACTCAACTCAGATGAGTACTGACAGCACTGGCTGAGGCTGAGTTG 854
 QY 841 CAGTGTGCTTGAAGGCGGAGTCAAGAGATGACTT 877
 Db 855 CAGTGTGCTTGAAGGCGGAGTCAAGAGATGACTT 891

RESULT 14

LOCUS BG708016 871 bp mRNA linear EST 07-MAY-2001
 DEFINITION 60267134f1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4793770 5',

mRNA sequence.

ACCESSION BG708016

VERSION BG708016

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 871)
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LLAM10674 row: b column: 11
High quality sequence stop: 839.

FEATURES

source

1. 871
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4793770"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 238 a 190 c 235 g 207 t 1 others
ORIGIN

Query Match 26.2%; Score 671; DB 12; Length 871;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 721; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 908 GGATGGTGGCAGGCCACATGATCCTGATGATGGGGAGAGACTTAACCACTGGGTTA 967
DB 5 GGATGGTGGCAGGCCACATGATCCTGATGATGGGGAGAGACTTAACCACTGGGTTA 64
QY 968 TAAGAGATATCCAAACGTGTTAAGAGATCCGAGCATTTGGAAGAGAGCGTACTGG 1027
DB 65 TAAGAGATATCCAAACGTGTTAAGAGATCCGAGCATTTGGAAGAGAGCGTACTGG 124
QY 1028 TGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAGAGCTGTGTTCCGGCATGAACGT 1087
DB 125 TGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAGAGCTGTGTTCCGGCATGAACGT 184
QY 1088 CAATGATCTGTACCAACAGAGATTGATACTGTACTGCTGCCGAGAAATCCATTTT 1147
DB 185 CAATGATCTGTACCAACAGAGATTGATACTGTACTGCTGCCGAGAAATCCATTTT 244
QY 1148 GGATGGCTGAAGAGAGACACAGATGTATGTTGGTGGGAACAAGTGGTGTGG 1207
DB 245 GGATGGCTGAAGAGAGACACAGATGTATGTTGGTGGGAACAAGTGGTGTGG 304
QY 1208 CTATGGTGAAGTAGGCAAGGGCTGTGCTGCTCTCAAGCTCTTGAGCAATTTGCTA 1267
DB 305 CTATGGTGAAGTAGGCAAGGGCTGTGCTGCTCTCAAGCTCTTGAGCAATTTGCTA 364
QY 1268 CATTACCGAAATGACCCCATCTGTGCTGTCAGGCTGCATGGATGGTTCAGGGTGT 1327
DB 365 CATTACCGAAATGACCCCATCTGTGCTGTCAGGCTGCATGGATGGTTCAGGGTGT 424
QY 1328 AAAGCTAAATGAAGTATCCGCAAGTGTGCTGATACTGACAGAAATAAGAA 1387
DB 425 AAAGCTAAATGAAGTATCCGCAAGTGTGCTGATACTGACAGAAATAAGAA 484
QY 1388 TGTAGTGACAGCGGAGCACTTGATCGCATGAATAACAGTGTATCGTATGCAATATGG 1447

DB 485 TGTAGTGACAGCGGAGCACTTGATCGCATGAATAACAGTGTATCGTATGCAATATGG 544
QY 1448 CCACCTCCACACAGAAATCGATGTGACACAGCCTCCGCACTCCGGAGCTGACGTGGAGCG 1507
DB 545 CCACCTCCACACAGAAATCGATGTGACACAGCCTCCGCACTCCGGAGCTGACGTGGAGCG 604
QY 1508 AGTACGTTCTCAGGTGAGACCATGTCATCTGGCCAGATGGCAACAGAGTTGCTCTCTGGC 1567
DB 605 AGTACGTTCTCAGGTGAGACCATGTCATCTGGCCAGATGGCAACAGAGTTGCTCTCTGGC 664
QY 1568 AGAGGTCGTCTACTCAATTGAGCTGCTCCACAGTTCCACCTTTGTTCTGTCATCAC 1627
DB 665 AGAGGTCGTCTACTCAATTGAGCTGCTCCACAGTTCCACCTTTGTTCTGTCATCAC 724
QY 1628 AG 1629
DB 725 AG 726

RESULT 15
BQ889173 933 bp mRNA linear EST 16-AUG-2002
LOCUS
DEFINITION AGENCOURT_8482604 lupski_dorsal_root_ganglion Homo sapiens cDNA
clone IMAGE:6185747 5', mRNA sequence.
ACCESSION BQ889173
VERSION BQ889173.1 GI:22281187
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 933)
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LLAM13577 row: a column: 12
High quality sequence stop: 656.

FEATURES

source

1. 933
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6185747"
/clone_lib="lupski_dorsal_root_ganglion"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGCGTCCG-3' and
5'-GACTAGTCTAGATCGGAGCGGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."

BASE COUNT 231 a 202 c 281 g 218 t 1 others
ORIGIN

Query Match 25.6%; Score 656; DB 14; Length 933;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 656; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 661 GTGCTCAGGGGAGAGCCCTTGCTGTGCTAAATAGTGGCTGTACACATCAG 720
|||||
Db 1 GTGCTCAGGGGAGAGCCCTTGCTGTGCTAAATAGTGGCTGTACACATCAG 60
QY 721 CCCAGACAGCGGTGTGATGAGACACTCTGTGCCCTGGGGCTCAGTGCCGCTGTCTG 780
|||||
Db 61 CCCAGACAGCGGTGTGATGAGACACTCTGTGCCCTGGGGCTCAGTGCCGCTGTCTG 120
QY 781 CTGTACATCTACTCACTCAGATGAGATGAGTGTGACACTGGCTGAGGCTGAGCTTG 840
|||||
Db 121 CTGTACATCTACTCACTCAGATGAGATGAGTGTGACACTGGCTGAGGCTGAGCTTG 180
QY 841 CAGTGTGCTGTGAGAGGGCGAGTCAAGAGATGACTTCTGTGTATTGACCGCTGTG 900
|||||
Db 181 CAGTGTGCTGTGAGAGGGCGAGTCAAGAGATGACTTCTGTGTATTGACCGCTGTG 240
QY 901 TGAACATGATGGGTGCGCAGGCCAACATGATCCTGATGATGGGGAGACTTAACCACT 960
|||||
Db 241 TGAACATGATGGGTGCGCAGGCCAACATGATCCTGATGATGGGGAGACTTAACCACT 300
QY 961 GGGTTTATAGAGATCCAAACGTGTTTAAGAGATCCGAGGCATTGTGGAAGAGAGCG 1020
|||||
Db 301 GGGTTTATAGAGATCCAAACGTGTTTAAGAGATCCGAGGCATTGTGGAAGAGAGCG 360
QY 1021 TGACTGGTGTACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTGTCCGCCA 1080
|||||
Db 361 TGACTGGTGTACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTGTCCGCCA 420
QY 1081 TGAACGTCAATGATCTGTACCACAGAGTTGATACTGTACTGTGCCGAGAAAT 1140
|||||
Db 421 TGAACGTCAATGATCTGTACCACAGAGTTGATACTGTACTGTGCCGAGAAAT 480
QY 1141 CCATTTTGATGGCTGAAGAGACACAGATGATGTTTGTGGGAACAAGTGTG 1200
|||||
Db 481 CCATTTTGATGGCTGAAGAGACACAGATGATGTTTGTGGGAACAAGTGTG 540
QY 1201 TGTGGCTATGTGAGGTAGGCAAGGGCTGTGTCTCTCAAAAGCTTTGAGCAA 1260
|||||
Db 541 TGTGGCTATGTGAGGTAGGCAAGGGCTGTGTCTCTCAAAAGCTTTGAGCAA 600
QY 1261 TTGCTACATTACCGAAATCGACCCCATCTGTCTGTGAGGCCCTGCATGATGG 1316
|||||
Db 601 TTGCTACATTACCGAAATCGACCCCATCTGTCTGTGAGGCCCTGCATGATGG 656

Search completed: April 20, 2003, 22:06:43
Job time : 3278 secs

Db 121 CGCAGCTGCAGCAGGCGCGCGAGAGGGTGGCGATCGCGTGTGAGGGCGCGCGC 180
QY 181 GGGCAGCGCGCGCGCGCGCGAGAGGGGAAAGAGCGCGCGCGCGGTCAAGCGCTGGCC 240
Db 181 GGGCAGCGCGCGCGCGCGCGAGAGGGGAAAGAGCGCGCGCGCGGTCAAGCGCTGGCC 240
QY 241 GGGCGCGCGCGCGCGCGCGATGTCGATGCTGACCGCGATGCCGTGCCCGCGCGCGAGAGC 300
Db 241 GGGCGCGCGCGCGCGCGCGATGTCGATGCTGACCGCGATGCCGTGCCCGCGCGAGAGC 300
QY 301 TGAAGCAGCGCGCGAGAGATCGAGAGCGCGAGAGTACTCCTTCATGCGCCACCGTCACCA 360
Db 301 TGAAGCAGCGCGCGAGAGATCGAGAGCGCGAGAGTACTCCTTCATGCGCCACCGTCACCA 360
QY 361 AGGCG 420
Db 361 AGGCG 420
QY 421 CCAAACTGGCG 480
Db 421 CCAAACTGGCG 480
QY 481 GTTCAGCTGCATCTCAGACAGATAGCTGATGATGAGGTTCTCCCGAGAGAGAGCAGC 540
Db 481 GTTCAGCTGCATCTCAGACAGATAGCTGATGATGAGGTTCTCCCGAGAGAGAGCAGC 540
QY 541 AAACCACTCCAGCG 600
Db 541 AAACCACTCCAGCG 600
QY 601 GACGCGCGCGAGATTGAGATTGACAGAGCAGACATGCTGCTGATTCATCAGGAAAC 660
Db 601 GACGCGCGCGAGATTGAGATTGACAGAGCAGACATGCTGCTGATTCATCAGGAAAC 660
QY 661 GTGCTCAGGGGGAGAGAGCG 720
Db 661 GTGCTCAGGGGGAGAGAGCG 720
QY 721 CCCAGACGCGGTGTGATTGAGACACTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
Db 721 CCCAGACGCGGTGTGATTGAGACACTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
QY 781 CTGTACATCTACTCACTCAGATGAGTAGCTGCAGCACTGGCTGAGCGCTGAGTTG 840
Db 781 CTGTACATCTACTCACTCAGATGAGTAGCTGCAGCACTGGCTGAGCGCTGAGTTG 840
QY 841 CAGTGTGCTTGAAGGGCGAGTCAAGATGACTTCTGCTGTATGACCGCTGTG 900
Db 841 CAGTGTGCTTGAAGGGCGAGTCAAGATGACTTCTGCTGTATGACCGCTGTG 900
QY 901 TGAACATGAGTGGGTGGCAGGCGCAACATGATCTGATGATGGGGGAGACTTAACCCACT 960
Db 901 TGAACATGAGTGGGTGGCAGGCGCAACATGATCTGATGATGGGGGAGACTTAACCCACT 960
QY 961 GGGTTATAGAAGTATCCAAACGTTTAAAGAGATCCGAGGCAATTGGAAGAGAGCG 1020
Db 961 GGGTTATAGAAGTATCCAAACGTTTAAAGAGATCCGAGGCAATTGGAAGAGAGCG 1020
QY 1021 TGACTGTGTTCAAGGCTGATCAGCTCTCCAAAGCTGGGAAGCTGTGTTCGGCCA 1080
Db 1021 TGACTGTGTTCAAGGCTGATCAGCTCTCCAAAGCTGGGAAGCTGTGTTCGGCCA 1080
QY 1081 TGAACGTCATGATCTGTTACCAACAGAGTTGATTAACCTGTACTGCTCCGAGAAAT 1140
Db 1081 TGAACGTCATGATCTGTTACCAACAGAGTTGATTAACCTGTACTGCTCCGAGAAAT 1140
QY 1141 CCATTTGATGCGCTGAAGAGGACAGATGATGTTGGTGGGAAACAAGTGTGG 1200
Db 1141 CCATTTGATGCGCTGAAGAGGACAGATGATGTTGGTGGGAAACAAGTGTGG 1200
QY 1201 TGTGTGCTATGTGAGTAGGCAAGGCTGCTGTGCTCAAAAGCTCTTGAGCAA 1260
Db 1201 TGTGTGCTATGTGAGTAGGCAAGGCTGCTGTGCTCTCAAAAGCTCTTGAGCAA 1260

Db 1201 TGTGTGCTATGTGAGTAGGCAAGGCGCTGCTGTGCTCTCAAAAGCTCTTGAGCAA 1260
QY 1261 TTGTCTACATTACCGAAATCGAACCCCATCTGTGCTGCGAGCGCTGATGATGGTTCA 1320
Db 1261 TTGTCTACATTACCGAAATCGAACCCCATCTGTGCTGCGAGCGCTGATGATGGTTCA 1320
QY 1321 GGGTGTAAAGCTAAATGAAGTCATCCGCGCAAGTCGATGCTGTAATACTTGACAGGAA 1380
Db 1321 GGGTGTAAAGCTAAATGAAGTCATCCGCGCAAGTCGATGCTGTAATACTTGACAGGAA 1380
QY 1381 ATAAGATGTAGTACACGGGAGCAGCAGTGGATCGCATGCCACCTCCGAGCTGACGT 1440
Db 1381 ATAAGATGTAGTACACGGGAGCAGCAGTGGATCGCATGCCACCTCCGAGCTGACGT 1440
QY 1441 ATATGGCCACTCCAAACAGAAATCGATGTGACAGCGCTCCGAGCTGACGT 1500
Db 1441 ATATGGCCACTCCAAACAGAAATCGATGTGACAGCGCTCCGAGCTGACGT 1500
QY 1501 GGGAGCGAGTACGTTCTCAGGTGGACCATGTGATCTGCGCCAGATGGCAACGAGTTGCC 1560
Db 1501 GGGAGCGAGTACGTTCTCAGGTGGACCATGTGATCTGCGCCAGATGGCAACGAGTTGCC 1560
QY 1561 TCCTGGCAGAGGGTCTCTACTCAATTGAGCTGCTCCACAGTTCCACCTTGTGTTCT 1620
Db 1561 TCCTGGCAGAGGGTCTCTACTCAATTGAGCTGCTCCACAGTTCCACCTTGTGTTCT 1620
QY 1621 CCATCAGCGCCAAACAGAGCGCTTGGCAGCTGATAGACTCTAATATGACCCGAGGCG 1680
Db 1621 CCATCAGCGCCAAACAGAGCGCTTGGCAGCTGATAGACTCTAATATGACCCGAGGCG 1680
QY 1681 GATACAGCAGAGATGTACTTGTCTTCTTAAGAAATGATGATAGCTTGCAGCTTGC 1740
Db 1681 GATACAGCAGAGATGTACTTGTCTTCTTAAGAAATGATGATAGCTTGCAGCTTGC 1740
QY 1741 ATCTGCCATCATTTGATGCGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTG 1800
Db 1741 ATCTGCCATCATTTGATGCGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTG 1800
QY 1801 GACTCAACAAATGGGCCATTCAACCTAATTATACAGATACTAATGACCATTACTAC 1860
Db 1801 GACTCAACAAATGGGCCATTCAACCTAATTATACAGATACTAATGACCATTACTAC 1860
QY 1861 CAAGGACAGTCCACCTGACACACACACTCTAAGAGAAATTTTAAAGATACTTTAT 1920
Db 1861 CAAGGACAGTCCACCTGACACACACACTCTAAGAGAAATTTTAAAGATACTTTAT 1920
QY 1921 TTCTCTTACTCTCTCTCTCTGATTTTCTCTAATTTCTATCTCTGTTTCTATC 1980
Db 1921 TTCTCTTACTCTCTCTCTCTGATTTTCTCTAATTTCTATCTCTGTTTCTATC 1980
QY 1981 TCATTATCCAAGTTCTGACAGACACACAGGAACTTGTCTATGAGTGAATA 2040
Db 1981 TCATTATCCAAGTTCTGACAGACACACAGGAACTTGTCTATGAGTGAATA 2040
QY 2041 GAAGTCAAGGTCCTCTACTCTAGTCACTAAGAGATTTTACTCCCCAGCCAGCAA 2100
Db 2041 GAAGTCAAGGTCCTCTACTCTAGTCACTAAGAGATTTTACTCCCCAGCCAGCAA 2100
QY 2101 GGTGATCTCTCTTACCATTTCTGGGACTTATGCTAATTAGGTACCTTATTAACA 2160
Db 2101 GGTGATCTCTCTTACCATTTCTGGGACTTATGCTAATTAGGTACCTTATTAACA 2160
QY 2161 GGAATGCTAAGGTAAGTCTCTGTTGAACAATCTGCAATGCTAATGCTTAAGA 2220
Db 2161 GGAATGCTAAGGTAAGTCTCTGTTGAACAATCTGCAATGCTAATGCTTAAGA 2220
QY 2221 GCCCATTTCTTACGCTGCAATCAGTGTCTTCTTCACTTCTCAGAGAGCAGGAGTGT 2280
Db 2221 GCCCATTTCTTACGCTGCAATCAGTGTCTTCTTCACTTCTCAGAGAGCAGGAGTGT 2280
QY 2281 ACCTACCGGCGAGGAGTATGATGTTGGTGTGATGTTAATTTCCCTTAGAGTTCCA 2340
Db 2281 ACCTACCGGCGAGGAGTATGATGTTGGTGTGATGTTAATTTCCCTTAGAGTTCCA 2340

QY	1	GGCGGGGCGAGGTCGGAGCTCGGAGCTGCTGCTTCTTGTTCTCTTGTGGCCACCGTCGCT	60
		:	
Db	4	GGCGCGGRCAGGTCGGAGCTCGAGCTGCTGCTTCTTGTTCTTGTGGCGCGCTGCT	63
QY	61	GTCGCGGCTGCTTGGGCTGCGCAACAGACGAGCGTGGGCACAGCACCTCAGAGCCGA	120
Db	64	GTCGCGGCTGCTTGGGCTGCGCAACAGACGAGCGTGGGCACAGCACCTCAGAGCCGA	123
QY	121	CGCAGCTCGACGCGAGGGGCGCGCAGAGGGTGGCGCATCGGTGTCGAGAGGCGCGCGC	180
Db	124	CGCAGCTCGACGCGAGGGGCGCGCAGAGGGTGGCGCATCGGTGTCGAGAGGCGCGCGC	183
QY	181	GGCGAGGCGGGGCGCGCCAGAGGGGGAAAGAGGGCGGGGCGGGTACGCCGTGGCC	240
Db	184	GGCGAGGCGGGGCGCGCCAGAGGGGGAAAGAGGGCGGGGCGGGTACGCCGTGGCC	243
QY	241	GGGCGCGCGGGGGAATGTCGATGCTTACCGCGATGCCGCTGCCCGGGGTGCGGGAGAGAC	300
Db	244	GGGCGCGCGGGGGAATGTCGATGCTTACCGCGATGCCGCTGCCCGGGGTGCGGGAGAGAC	303
QY	301	TGAAGCAGGCCAAGGAGATCGAGGACGCGCGAGAAGTACTCTTCATGGCCACCGTCACCA	360
Db	304	TGAAGCAGGCCAAGGAGATCGAGGACGCGCGAGAAGTACTCTTCATGGCCACCGTCACCA	363

OY	1322	GGTGGTAAAGCTAAATGAAGTCATCCGGCAAGTCGATGTGTAATAACTTGACAGGAAA	1381
Dp	1	GGTGGTAAAGCTAAATGAAGTCATCCGGCAAGTCGATGTGTAATAACTTGACAGGAAA	60
OY	1382	TAAAGAATGTAGTGACACACGGGAGACACTTGGATCGCATGA AAAACAAGTTGATCGTATGC AA	1441
Dp	61	TAAAGAATGTAGTGACACACGGGAGACACTTGGATCGCATGA AAAACAAGTTGATCGTATGC AA	120
OY	1442	TATGGGCCACTCCACACACAGAAATCGATGTGACCAGCCTCCGCAC TCCGGAGCTGACGTG	1501
Dp	121	TATGGGCCACTCCACACACAGAAATCGATGTGACCAGCCTCCGCAC TCCGGAGCTGACGTG	180
OY	1502	GGAAGCGAGTAGCTTCTCAGAGGTGACCATGTTCATCTGGCCAGATGSCAAACGAGTTGT CCT	1561
Dp	181	GGAAGCGAGTAGCTTCTCAGAGGTGACCATGTTCATCTGGCCAGAT -GNAAACGAGTTGT CCT	239
OY	1562	CCATGGCAGAGGGTCTCTACTCAATTGAGCTGCTCCACAGATGCCACCTTGTGTCTG TC	1621
Dp	240	CCATGGCAGAGGGTCTCTACTCAATTGAGCTGCTCCACAGATGCCACCTTGTGTCTG TC	299

OY 1622 CATCAGCCACACACAGGCTTTGGCACTGATAGACTCTATATGACCCGAGGGCG 1681
DB 300 CATCAGCCACACACAGGCTTTGGCACTGATAGACTCTATATGACCCGAGGGCG 359
OY 1682 ATACAAGCAGAGTGTGACTTGTCTTCTTAAGAAA-TGGATGAATACGTTGCCAGCTTGC 1740
DB 360 ATACAAGCAGAGTGTGACTTGTCTTCTTAAGAAAATGGATGAATACGTTGCCAGCTTGC 419
OY 1741 ATCTGCATCATTTGATGCCCCACCTTACAGAGCTGACAGATGACCAAGAAAATATCTGG 1800
DB 420 ATCTGCATCATTTGATGCCCCACCTTACAGAGCTGACAGATGACCAAGCTGAGTCTGG 478
OY 1801 GACTCAACAAAATGGCCCATTCACACCTAATTTATACAGATACTAATGACCACTACTAC 1860
DB 479 GACTCAACAAA--TGGGCCATTCACACCTAATTTATACAGATACTAATGACCACTACTAC 536
OY 1861 CAAGGACCAGTCCAC 1875
DB 537 CAAGGACCAGTCCAC 551

RESULT 4

US-09-879-536-850/C
; Sequence 850, Application US/09879536
; Patent No. US20020144298A1
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dertl, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/879,536
; CURRENT FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/088,801
; PRIOR FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for windows version 3.0
; SEQ ID NO 850
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(636)
; OTHER INFORMATION: n = A,T,C or G
US-09-879-536-850

Query Match 15.9%; Score 406.4; DB 10; Length 636;
Best Local Similarity 99.5%; Pred. No. 3.6e-105;
Matches 418; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 707 TACACACATCACAGCCAGACAGC-GGTGTGATGAGACACTGTGCCCTGGGGGCTC 765
DB 420 TACACACATCACAGCCAGACAGCAGGGGTGTGATGAGACACTGTGCCCTGGGGGCTC 361
OY 766 AGTCCGCTGTCTGCTTGTACATCTACTCAACTCAGATGAGTAGCTGACGACACTGG 825
DB 360 AGTCCGCTGTCTGCTTGTACATCTACTCAACTCAGATGAGTAGCTGACGACACTGG 301
OY 826 CTGAGGCTGTGAGTTCAGTGTTCGCTTGGAGGGCGAGTCAGAGATGACTTCTGTGTGT 885
DB 300 CTGAGGCTGTGAGTTCAGTGTTCGCTTGGAGGGCGAGTCAGAGATGACTTCTGTGTGT 241

OY 886 GTATTGACCGCTGTGTGAACATGATGGGTGGCAGGCCAACATGATCTGTGATGGG 945
DB 240 GTATTGACCGCTGTGTGAACATGATGGGTGGCAGGCCAACATGATCTGTGATGGG 181
OY 946 GAGACTTAACCCACTGGGTTTATAAGAGTATCCAAACGTTTAAAGATCCGAGGA 1005
DB 180 GAGACTTAACCCACTGGGTTTATAAGAGTATCCAAACGTTTAAAGATCCGAGGA 121
OY 1006 TTGTGGAAGAGAGCGGTGACTGTCTTACAGAGCTGTATCAAGCTCTCCAAAGCTGGAGC 1065
DB 120 TTGTGGAAGAGAGCGGTGACTGTCTTACAGAGCTGTATCAAGCTCTCCAAAGCTGGAGC 61
OY 1066 TCTGTGTCCGGCCATGACGTCATGATTTCTGTACCACAGAGTTGATACTTGT 1125
DB 60 TCTGTGTCCGGCCATGACGTCATGATTTCTGTACCACAGAGTTGATACTTGT 1

RESULT 5

US-09-925-301-217
; Sequence 217, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05382
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 217
; LENGTH: 2200
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2188)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-217

Query Match 15.8%; Score 404; DB 10; Length 2200;
Best Local Similarity 58.4%; Pred. No. 3.9e-104;
Matches 745; Conservative 0; Mismatches 525; Indels 6; Gaps 2;

OY 580 ACATCAAGCAGCAGAAATTTGACGCCGCGAGATGAGATTGACAGCAAGACATGTCTG 639
DB 108 ACATCGGCGCTGTGCTGCGGAGCGCAAGGCCCTGACATGTCTGAGACAGATGCCG 167
OY 640 CTCTGATTTCACTCAGGAACGTGCTCAGGGGAGAGGCCCTGGCTGTCTAAATAG 699
DB 168 GCCTGATGCGTATGCGGAGCGGTACTCGGCCCTCCAAGCCACTGAAGGGCGCCGATCG 227
OY 700 TGGGCTGTACACACATCACAGCCAGACAGCGGTGTGATGAGACACTCTGTGCCCTGG 759
DB 228 CTGGCTGCTGCACATGACCGTGAAGACGGCGCTCCTCATTTGAGACCCCTGTCACCTGG 287
OY 760 GGGCTCAGTGGCGGTGCTGCTTGTACATCTACTCAACTCAGATGAGTAGCTGCAG 819
DB 288 GTGCTGAGGTGACGTGTGTCAGCTGCAACATCTTCTCAAGCCAGAGACATGCGGGCTG 347
OY 820 CACTGGCTGAGCGCTGAGTTCAGTGTTCGCTTGGAGGGCGAGTCAGAGATGACTTCT 879
DB 348 CCATTCGCAAGCGCTGCAATTCGGGTGTATGCCCTGGAAGGGCGAAGGAGAGTAC 407
OY 880 GGTGTGTATTGACCGCTGTGTGAACATGATGGGTGGAGGCCAACATGATCTGTGATG 939
DB 408 TGTGTGATTTGAGCAGACCCCTGTACTTCAAGAGAGGGGCCCTCAACATGATTTGTGAG 467
OY 940 ATGGGGAGACTTAACCCACTGGGTTTATAAGAGATATCCAAACGTTTAAAGAGATCC 999

DB 468 ACGGGGGGACCTTCAACCACTCATCCACACCAAGTACCCGAGCTTCTGCCAGGCAATCC 1111
QY 1000 GAGGCATGTGGAGAGAGCGGTGACTGGTGTTCACAGGCTGTATCAGCTCTCCAAAGCTG 1059
DB 528 GAGGCATCTGTGAGAGAGACCAAGCTGGGGGTCCACAACTCTTCAAGATGATGGCCAAATG 587
QY 1060 GGAAGCTCTGTGTCCGGCCATGAAAGCTCAATGATCTGTACCAACAGAGATTGATA 1119
DB 588 GGATCTCAAGGTGCTTCCATCAATGATGATCTCCATCAACCAAGAGCAAGTTGACA 647
QY 1120 ACTGTACTGTGCTGAGAAATCCATTTGGATGGCTGAGAGAGACCAAGATGTGATG 1179
DB 648 ACCTGTATGCTGCTGGAGAGTCCCTCATAGATGGCAACAGCGGCCACAGATGTGATGA 707
QY 1180 TTGGTGGAAACAGTGGT 1239
DB 708 TTGCGGCAAGGTAGCGGT 767
QY 1240 CTCTCAAGCTCTGAGCAATTTGTCTACATTTACCAAAATGACCCCATCTGTGCTGTC 1299
DB 768 CCTGCGGGGTTCGAGAGCCCGGCTCATCATCAACGAGATGACCCCATCAACGCACTGC 827
QY 1300 AGGCTGCATGATGGTTCAGGGTGTAAAGCTAAATGAAGTATCCGGCAAGTGCATG 1359
DB 828 AGGCTGCCATGAGGGCTATGAGGTGACCAACCATGATGAGGCCCTGTCAAGAGGGCAACA 887
QY 1360 TCGTAATTAATTCGACAGGAATAAGATGTAGTACACGGGAGACACTTGGATGCGATGA 1419
DB 888 TCTTTGTACACACAGAGCTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 947
QY 1420 AAAACAGTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1479
DB 948 AGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1007
QY 1480 TCCGCACTCCGAGCTGACGTGGGAGGAGTACGTTCAGGTGAGTGCATGTCGTC 1539
DB 1008 TCAACGAGAACGCGGTGAGAGAGTGAACATCAAGCCGAGTGGAGCGGTATGCGTGA 1067
QY 1540 CAGATGGCAAGAGTGTCTCTCTGCGAGAGGGTGTCTACTCAATTTGAGCTGTGCA 1599
DB 1068 AGAATGGGCGCGCATCATCTGCTGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1127
QY 1600 CAG--TTCCACACTTGTCTGTCTCATCACAGCACAACAGGCTTGGCACTGATAG 1656
DB 1128 TGGGCGACCCAGCTTCGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1187
QY 1657 AACTCTAATATGACCCGAGGGCGATACAAAGCAGATGTACTGCTCTCTAAGAAA 1716
DB 1188 AGCTGTGACCCATCAGACAAG--TACCCCGTGGGGTTCATTTCTGCCCCAAGAAC 1244
QY 1717 TGGATGAATACGTGGCAGCTTGCATCTGCCATCATTTGATGCCACCTTACAGAGCTGA 1776
DB 1245 TGGATGAGGAGTGGCTGAAGCCCACTGGGCAAGCTGAATGTGAAGTTGACCAAGCTAA 1304
QY 1777 CAGATGACCAAGCAAAATATCTGGGACTCAACAAAAATGGCCATTCAAACCTAATTA 1836
DB 1305 CTGAGAAGCAAGCCCACTACCTGGGCACTGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1364
QY 1837 ACAGATACTAATGAC 1852
DB 1365 ACAGTACTGAGAGCC 1380

RESULT 6
US-10-044-090-344

; Sequence 344, Application US/10044090

; Patent No. US20020137081A1

; GENERAL INFORMATION:

; APPLICANT: Olga Bandman

; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION

; FILE REFERENCE: PA-0028 US

; CURRENT APPLICATION NUMBER: US/10/044, 090

; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 344
; LENGTH: 2429
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 1468237CBI
US-10-044-090-344

Query Match 15.8%; Score 404; DB 12; Length 2429;
Best Local Similarity 58.4%; Pred. No. 4.2e-104;
Matches 745; Conservative 0; Mismatches 525; Indels 6; Gaps 2;

QY 580 ACATCAAGCAGCAGAAATTTGGACCGCGGAGATGAGATTGCAAGCAAGACATGTCTG 639
DB 108 ACATCGGCTGGCTGCTGCGGAGCGCAAGGCCCTGGACATTGCTGAGAAACGAGATGCCG 167
QY 640 CTCTGATTTCACTCAGGAACGCTGCTCAGGGGAGAGAGCCCTTGGCTGCTAAATAG 699
DB 168 GCGTATGCGTATGCGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 227
QY 700 TGGGCTGACACATCAGAGCCAGAGCGGTGTGATGAGACACTGTGCTGCTGCTGCTG 759
DB 228 CTGCTGCTGACATGACCGTGGAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 287
QY 760 GGGCTAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 819
DB 288 GTGCTAGTGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 347
QY 820 CACTGCTGAGGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 879
DB 348 CCATGCGCAAGGCTGCGCATTCGGGTGATGCTGGAAGGGCGAAGCGAAGAGAGTACC 407
QY 880 GGTGCTATTTGACCGCTGTGTGAACATGATGGGTGGCAGGCCAATGATCTGATG 939
DB 408 TGTGCTGATTTGAGCAGAACCTGTACTTCAAGGACGGGCCCTCAACATGATCTGAGC 467
QY 940 ATGGGGGAGCTTAACCCACTGGGTTTAAGAGTATCCAAACGCTGTTAAGAGATCC 999
DB 468 ACGGGGGAGCTTCAACCACTCATCAACCAAGTACCCGAGCTTCTGCGAGCATCC 527
QY 1000 GAGGCATGTGGAAGAGAGCGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1059
DB 528 GAGGCATCTGTGAGAGAGACCAAGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 587
QY 1060 GGAAGCTGTGTCTCCGCGCATGAAAGCTAATGATGCTGTTACCAACAGAGTTGATA 1119
DB 588 GGATCTCAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 647
QY 1120 ACTGTACTGTGCTGCGGAGAAATCCATTTGGATGGCTGGAAGAGGACACAGATGTGAT 1179
DB 648 ACCTGTATGCTGCGGAGAGTCCCTCATAGATGGCATCAAGCGGGCACAGATGTGATGA 707
QY 1180 TTGGTGGAAACAGTGGT 1239
DB 708 TTGCGGCAAGGTAGCGGT 767
QY 1240 CTCTCAAGCTCTTGGAGCAATTTGTCTACATTTACCAAAATGACCCCATCTGTGCTGTC 1299
DB 768 CCTGCGGGGTTCGAGAGCCCGGCTCATCATCAACGAGATGACCCCATCAACGCACTGC 827
QY 1300 AGGCTGCATGATGGTTCAGGGTGTAAAGCTAAATGAAGTATCCGGCAAGTGCATG 1359
DB 828 AGGCTGCCATGAGGGCTATGAGGTGACCAACCATGATGAGGCCCTGTCAAGAGGCAACA 887
QY 1360 TCGTAATTAATTCGACAGGAATAAGATGTAGTACACGGGAGACACTTGGATGCGATGA 1419
DB 888 TCTTTGTACACACAGAGCTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 947
QY 1420 AAAACAGTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1479

Db 948 AGGATGATGCCATTTGTGTACATTTGACACTTTGACGTGAGATGATTCACAGTGGC 1007
QY 1480 TCCGCACTCCGAGCTGACGTGGAGCGAGTACGTCTCTCAGGTGGACCATGTCATCTGGC 1539
Db 1008 TCAACGAGAACGCCGTGGAGAGGTGACATCAAGCCGAGGTGGACCGGTATCGGTGA 1067
QY 1540 CAGATGGCAACGAGTGTCTCTCTGCGAGAGGGTCTCTACTCAATTGAGCTGTCCA 1599
Db 1068 AGAATGGCGCCGATCATCTCTGCGAGGGTCTGCTGTCACACTGGGTGTGCCA 1127
QY 1600 CAG--TTCCCACTTTGTTCTGTCTCCATCAGCCACACACAGAGCTTGGCACTGATAG 1656
Db 1128 TGGGCCACCCAGCTTCGTGATGAGTAACTCTTCAACCAACAGGTGATGGCGAGATCG 1187
QY 1657 AACTCTAATATGCAACCCGAGGGCGGATACACAGAGTGTACTTGTCTTCTAAGAAA 1716
Db 1188 AGCTGTGACCCATCCAGACAG--TACCCCGTTGGGGTTCAATTCCTGCCCAAGAAAGC 1244
QY 1717 TGGATGAATAGCTTCCAGCTTGCATCTGCCATTCATTTGATGCCACCTTACAGAGCTGA 1776
Db 1245 TGGATGAGGAGCTGGCTGTAAGCCACCTGGGCAAGCTGAATGTGAAGTTGACCAAGCTAA 1304
QY 1777 CAGATGACCAAGCAAAATATCTGGGACTCAACAAATGGGCCATTCAACTAATTAAT 1836
Db 1305 CTGAGAAGCAAGCCAGTACCTGGGATGTCTGTATGGCCCTTCAAGCCGAGTCACT 1364
QY 1837 ACAGATACTAATGAC 1852
Db 1365 ACCGCTACTGAGAGCC 1380

RESULT 7

US-09-880-107-1031/c
; Sequence 1031, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1031
; LENGTH: 376
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA432168
US-09-880-107-1031

Query Match 14.2%; Score 365; DB 10; Length 376;
Best Local Similarity 99.7%; Pred. No. 1.5e-93;
Matches 376; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 2174 TACCTTCTCTGGAACAATCTGCAATGCTAAATCGCCTTAAAGAGCCCATTTCTTAG 2233
Db 376 TACCTTCTCTGGAACAATCTGCAATGCTAAATCGCCTTAAAGAGCCCATTTCTTAG 317
QY 2234 CTGCTGAATCAGTGTCTTCTTCACTCTCTGAGAGAGCAGGATGTACCTACCCGCGAG 2293
Db 316 CTGCTGAATCAGTGTCTTCTTCACTCTCTGAGAGAGCAGGATGTACCTACCCGCGAG 257
QY 2294 GTAGTTAGATGTGGTGTGTCATGTAAATTTCCCTTAGAAGTTCACAGCCCTGTTTCT 2353
Db 2353 GTAGTTAGATGTGGTGTGTCATGTAAATTTCCCTTAGAAGTTCACAGCCCTGTTTCT 2353

Db 256 GTAGTTAGATGTGGTGTGTCATGTAAATTTCCCTTAGAAGTTCACAGCCCTGTTTCT 197
QY 2354 GCGTAAAGGTGTATGTCAGTTCAGAGATGTATATAGCATGCGTTGTTAGATCA 2413
Db 196 GCGTAAAGGTGTATGTCAGTTCAGAGATGTATATAGCATGCGTTGTTAGATCA 137
QY 2414 GGAGCCCACTTGTGATTTATAGTATAGCCCTTCCCTCCACTCCCAACAGACTTGTCTATT 2473
Db 136 GGAGCCCACTTGTGATTTATAGTATAGCCCTTCCCTCCACTCCCAACAGACTTGTCTATT 77
QY 2474 TTGAGTTTAACTAGACTACACTTATTGAGTTTAAATTTTGTCTCTAGATTATT 2533
Db 76 TTGAGTTTAACTAGACTACACTTA-TTGAGTTTAAATTTTGTCTCTAGATTATT 18
QY 2534 TCTGTTGCCAAAAA 2550
Db 17 TCTGTTGCCAAAAA 1

RESULT 8.

US-09-880-107-346
; Sequence 346, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 346
; LENGTH: 384
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA157401
US-09-880-107-346

Query Match 13.1%; Score 337; DB 10; Length 384;
Best Local Similarity 99.0%; Pred. No. 1.4e-85;
Matches 381; Conservative 0; Mismatches 0; Indels 4; Gaps 4;
QY 992 GAAGATCCGAGCATTTGTGAGAGAGAGCGTACTGTTGTCACAGGCTGTATCAGCTCTC 1051
Db 1 GAAGATCCGAGCATTTGTGAGAGAGAGCGTACTGTTGTCACAGGCTGTATCAGCTCTC 60
QY 1052 CAAAGCTGGAGAGCTCTGTCTCCGGCCATGACGTCAATGATTTCTTACCAACAGAA 1111
Db 61 CAAAGCTGGAGAGCTCTGTCTCCGGCCATGACGTCAATGATTTCTTACCAACAGAA 119
QY 1112 GTTGATACTTGTACTGCTGCCGAGATCCATTTTGGATGGCTGAAGAGACACAGA 1171
Db 120 GTTGATACTTGTACTGCTGCCGAGATCCATTTTGGATGGCTGAAGAGACACAGA 179
QY 1172 TGTGATGTTGTGGGAACAAGTGTGTGTGTGGCTATGCTGAGGTAGGCAAGGCTG 1231
Db 180 TGTGATGTTGTGGGAACAAGTGTGTGTGTGGCTATGCTGAGGTAGGCAAGGCTG 239
QY 1232 -CTGTGCTGCTCTC-AAAGCTCTTGAGCAATTTGTCTACATTAACGAAATGACCCCATC 1289
Db 240 CTTGTGCTGCTCTCAAAAGCTCTTGAGCAATTTGTCTACATTAACGAAATGACCCCATC 299
QY 1290 TGTGCTCTGAGGCTGATGATGAGTGTGAGGCTGTT-AAAAGTAAATGAGTCAATCG 1348
Db 300 TGTGCTCTGAGGCTGATGATGAGTGTGAGGCTGTT-AAAAGTAAATGAGTCAATCG 359

OY 1349 GCAAGTCGATGTCGTATACTGC 1373
|||||
Db 360 GCAAGTCGATGTCGTATACTGC 384

RESULT 9
US-09-884-441-40/c

; Sequence 40, Application US/09884441
; Patent No. US20020119158A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C7
; CURRENT APPLICATION NUMBER: US/09/884,441
; CURRENT FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 40
; LENGTH: 292
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-884-441-40

Query Match 10.2%; Score 261.4; DB 10; Length 292;
Best Local Similarity 99.3%; Pred. No. 3.9e-64;
Matches 273; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 2261 TTCAGAGAAGCAGGATGTACTACCCGGCAGGTAGATGTGGTGCATGTT 2320
|||||
Db 282 TTCAGCAGGAGGATGTACTACCCGGCAGGTAGATGTGGTGCATGTT 223
OY 2321 AATTCCCTAGAGTCCAGCCCTGTTCCGCGTAAGGTGATGTCCAGTTCAGA 2380
|||||
Db 222 AATTCCCTAGAGTCCAGCCCTGTTCCGCGTAAGGTGATGTCCAGTTCAGA 163
OY 2381 GATGTATATGAGCATGGCTGTGTAAGATCAGAGGCCCACTTGATTTATAGTAG 2440
|||||
Db 162 GATGTATATGAGCATGGCTGTGTAAGATCAGAGGCCCACTTGATTTATAGTAG 103
OY 2441 CCCTCCCTCCACTCCACAGACTGCTCATTTTTCAGAGTTTAACTAGACTCACTCT 2500
|||||
Db 102 CCCTCCCTCCACTCCACAGACTGCTCATTTTTCAGAGTTTAACTAGACTCACTCT 43
OY 2501 ATTTGAGTTTAAATTTGCTCTAGAGTTTATTC 2535
|
Db 42 A-TTGAGTTTAAATTTGCTCTAGAGTTTATTC 9

RESULT 10

US-09-759-990-1
; Sequence 1, Application US/09759990
; Patent No. US20020119491A1
; GENERAL INFORMATION:
; APPLICANT: Anticancer, Inc.
; APPLICANT: Xu, Mingxu
; APPLICANT: Han, Qianhong
; TITLE OF INVENTION: HIGH EXPRESSION AND PRODUCTION OF HIGH
; TITLE OF INVENTION: SPECIFIC ACTIVITY RECOMBINANT S-ADENOSYLHOMOCYSTEINASE
; TITLE OF INVENTION: (SAHH) AND IMPROVED ASSAYS FOR S-ADENOSYLMETHIONINE (SAM)
; FILE REFERENCE: 31276-20026.00
; CURRENT APPLICATION NUMBER: US/09/759,990
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/176,444
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1461
; TYPE: DNA
; ORGANISM: Unknown

; FEATURE:
; OTHER INFORMATION: Nucleotide sequence encoding SAHH
US-09-759-990-1

Query Match 9.5%; Score 244; DB 10; Length 1461;
Best Local Similarity 57.9%; Pred. No. 1e-58;
Matches 493; Conservative 0; Mismatches 350; Indels 9; Gaps 3;

OY 1002 GGCATGTGGAAGAGAGCGGTGACTGTTGTCACAGCGCTGTATCAGCTCTCCAAGCTGG 1061
|||
Db 613 GGTGTTCCGAAGAGACAACAGAGGTGTCACCGCCTCTACAGCTCGAAGAGAGGC 672
OY 1062 AAGCTCTGTGTTCCGGCCAGTACGTCATGATTTCTGTACCAACAAGATTGATTAAC 1121
|||
Db 673 AAGCTCTCTTCCAGCCATCAAGCTCAACGACGCTGTACCAAGTTCAGATTAAC 732
OY 1122 TTGTAAGTCTGCCGAGAAATCCATTTGGATGGCCTGAAGAGACACAGATGTGATTT 1181
|||
Db 733 ATCTACGGCTGCCGCCACTCCCTATCGATGTATCAACCGTCTCCGATGTATGATTC 792
OY 1182 GGTGGAAACAAGTGTGTTGCTGCTATGTGAGGTAGGCAAGGCTGTGCTGCT 1241
|||
Db 793 GCGGCAAGACAGCTCTGCTATGGGTTACGGCATGTCCGCAAGGCTGCGCTCAATCC 852
OY 1242 CTCAAGCTCTTGAGCAATGCTTACATTAACGAAATCGACCCATCTGCTCTGCAG 1301
|||
Db 853 CTGCGTGGCCAAAGGCGCTGCTGATCATCAAGAACTGACCCCAATCTGCGCTCCAG 912
OY 1302 GCTTCATGATGGTTCAGGTTAAAGCTAAATGAAGTCAATCCGCAAGTGCATGTC 1361
|||
Db 913 GCTGCATGGAAGGCTACAGAGTCCGCCCATCGAAGAGTCTCAAGATGTGATATC 972
OY 1362 GTAATTAAGTTCACAGAAATGAATGATGACACAGGAGCACTTGATCGCATGAA 1421
|||
Db 973 TGTGATACATGACAGAAACTGCAATATCATCTCTGTGACATGATGGCCAGATGAA 1032
OY 1422 AACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1481
|||
Db 1033 GATAGGCTATGTCGTAACATGCGCCACTTCGATACGAATGATGATGATGATGAT 1092
OY 1482 CGCA---CTCCGAGCTGACGTGGAGCGAGTACGTTCTCAGGTGACCATGTCATCTG 1538
|||
Db 1093 ATGAATATCCAGGATCAAGCATCCCAATCAAGCCAGAAATGACATGTGGGAATTC 1152
OY 1539 CCAGATGGCAAGAGTGTCTCTCTGAGAGGCTGCTACTCAATTTGAGCTGCTCC 1598
|||
Db 1153 CCAGATGGCAAGAGTGTCTCTCTGAGAGGCTGCTACTCAATTTGAGCTGCTCC 1212
OY 1599 ACAAT---CCACCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1655
|||
Db 1213 ACAGTCAACCATCTTCT 1272
OY 1656 GAATCTATATGACACCCGAGGCGCATACAGCAGAGTGTGATCTGCTCTAGAA 1715
|||
Db 1273 GACTCTTACGA---AAAGAGAGAAATCTCGAAGAAAGTTTACACACTTCCGAAGCAT 1329
OY 1716 ATGATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1775
|||
Db 1330 CTGATGAAGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1389
OY 1776 ACAGATGACCAAGCAAAATATCTGAGACTCAACAAATGAGGCTTCAAACTAATAT 1835
|||
Db 1390 ACAGAGAGCAGGCTGACTACATCAAGTTCAGTTGAGGCTCTTCAAGTCTGATGCT 1449
OY 1836 TACAGATACTAA 1847
|||
Db 1450 TACGTTATTTAA 1461

RESULT 11

US-09-746-660A-97
; Sequence 97, Application US/09746660A
; Publication No. US20030049804A1

```

: GENERAL INFORMATION:
: APPLICANT: Pompejus, Markus
: APPLICANT: Krieger, Burkhard
: APPLICANT: Schroder, Hartwig
: APPLICANT: Zeider, Oskar
: APPLICANT: Haberhauer, Gregor
: APPLICANT: Klm, Jun-Won
: APPLICANT: Lee, Heung-Schick
: APPLICANT: Hwang, Byung-Joon
: TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
: TITLE OF INVENTION: METABOLIC PATHWAY PROTEINS
: FILE REFERENCE: BGI-121CP2
: CURRENT APPLICATION NUMBER: US/09/746,660A
: CURRENT FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 09/606740
: PRIOR FILING DATE: 2000-06-23
: PRIOR APPLICATION NUMBER: 09/603124
: PRIOR FILING DATE: 2000-06-23
: PRIOR APPLICATION NUMBER: 60/141031
: PRIOR FILING DATE: 1999-06-25
: PRIOR APPLICATION NUMBER: 60/142101
: PRIOR FILING DATE: 1999-07-02
: PRIOR APPLICATION NUMBER: 60/148613
: PRIOR FILING DATE: 1999-08-12
: PRIOR APPLICATION NUMBER: 60/187970
: PRIOR FILING DATE: 2000-03-09
: PRIOR APPLICATION NUMBER: DE 19931420.9
: PRIOR FILING DATE: 1999-07-08
: NUMBER OF SEQ ID NOS: 125
: SOFTWARE: Patentln Vers. 2.0
: SEQ ID NO 97
: LENGTH: 1557
: TYPE: DNA
: ORGANISM: Corynebacterium glutamicum
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (101)..(1534)
: OTHER INFORMATION: RXN00132
: US-09-746-660A-97

Query Match          9.0%; Score 230.6; DB 9; Length 1557;
Best Local Similarity 57.0%; Pred. No. 6.9e-55;
Matches 482; Conservative 0; Mismatches 354; Indels 9; Gaps 3;

QY 1011 GAAGAGAGCGTGACTGCTGTTACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGT 1070
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 698 GAGGAAACCAACACCGGTGTGACCGGCTGTACCACTTCCGTGAAGAGGCGTGTGCTT 757
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 1071 GTTCCGCCATGAACGTCATGATTTCTGTTACCAACAGAGTTGATTAATCTGTACTGC 1130
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 758 TTCCACGCGATGAACGTCACAGCGCTGTACCAAGTCCAAAGTTGATTAACAAGTACGGC 817
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 1131 TGGCGAGATCCATTTTGGATGCGCTGAAGAGGACACAGATGATGTTGGTGGAAA 1190
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 818 ACCCGCACTCCCTGATCGAGCGGATCAACCGCGCACTGACATGCTCATGGCGGCAAG 877
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 1191 CAAGTGTGTGTGTGCTATGCTGAGGTAGGCAAGGCGTGTGCTGCTCTCAAGCT 1250
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 878 AACGTGCTGTGTGCGGTACGCGATGTCGCAAGGCGTGTGAGGCTTTCGACGGC 937
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 1251 CTGGAGCAATTTGTACATTAACGAAATGACCCCATCTGTGCTCTGACGCGCTGCATG 1310
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 938 CAGGCGGCTCGGCTCAAGGTCACGGAAGCTGACCCCAATCAACGCTCTTCAGGCTCTGATG 997
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 1311 GATGGGTTCAAGGTGTAAGCTAAATGAAGTATCCCGCAAGTGCATGTCGTAATACT 1370
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 998 GATGGCTACTCTGTGTACCGCTGATGAGGCCATGAGGAGCGCGGACATCTGTGATCACC 1057
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 1371 TGCACAGGAATAAGATGATGACACGCGGAGCAGTGTGATCGCATGAAAAACAGTTGT 1430
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1058 GCGACCGGCAACAGGACATCAATTCCTTGAGCAGATGCTCAAGATGAAGATCAGCT 1117
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 1431 ATCGTATGCAATATGAGGCACTCCAAACAGAAATGATGTG---ACCAGCCTCCGCACT 1487
```

```

    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1118 CTGCTGGCAACATCGGTCACTTTGATATAGATGATGATTCATTCCTGTGACCGC 1177
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 1488 CCGAGCTGACGTGGAGCGAGTACGTTCTCAGGTGAGCCATGTCATCTGGCAGATGGC 1547
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1178 GACGACGTACCCCGACACGATCAAGCCACAGGTGACAGATTCACCTTCTCCACGGT 1237
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 1548 AAAGAGTTGTCTCTCGGAGAGGCTGTCTACTCAATTTGAGCTGCTCCAC--AGTT 1604
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1238 CGCTCCATCATCGTCTGTCGGAAGGTGCGCTGTGTAACCTTGGCAACGCCACCGGAC 1297
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 1605 CCCACCTTGTCTCTCAATCAAGCCACACACAGGCTTTGGCACTGATAGACTCTAT 1664
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1298 CCATCATTTGTCAATGTCACACTCTTTCGCCGATCAGACATTTGCCAGATCGAATGTT 1357
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 1665 AATGACCCGAGGCGGATACAGCAGATGTGTACTGCTTCTTAAGAAATGATGAA 1724
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1358 CA--AAACGAAGACAGTACGAGAAGAGTGTACCGTGTGCTTAAGGTTCTGACGAA 1414
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 1725 TACGTGCGAGCTTGCAATCTGCCATCATTTGATGCCCACTTACAGAGCTGACAGATGAC 1784
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1415 AAGGTGACGACATCCAGCTTGAGGCTCTCGCGGTCACTCACCAGACTGACCAAGAG 1474
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 1785 CAAGCAAAATATCTGGGACTCAACAAAATGGGCCATCAAACTAATTATACAGATAC 1844
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1475 CAGCTGAGTACATCGCGGTGACGTTGACAGGCCCATTCAGCCGAGACACTACCGCTAC 1534
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 1845 TAATG 1849
    |||||
Db 1535 TAATG 1539
```

```

RESULT 12
US-09-738-626-836
: Sequence 836, Application US/09738626
: Publication No. US20020197605A1
: GENERAL INFORMATION:
: APPLICANT: NAKAGAWA, SATOSHI
: APPLICANT: MIZOGUCHI, HIROSHI
: APPLICANT: ANDO, SEIKO
: APPLICANT: HAYASHI, MIKIRO
: APPLICANT: OCHIAI, KEIKO
: APPLICANT: YOKOI, HARUHIKO
: APPLICANT: TATEISHI, NAKO
: APPLICANT: SENOH, AKIHIRO
: APPLICANT: IKEDA, MASATO
: APPLICANT: OZAKI, AKIO
: TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
: FILE REFERENCE: 249-125
: CURRENT APPLICATION NUMBER: US/09/738,626
: CURRENT FILING DATE: 2000-12-18
: PRIOR APPLICATION NUMBER: JP 99/377484
: PRIOR FILING DATE: 1999-12-16
: PRIOR APPLICATION NUMBER: JP 00/159162
: PRIOR FILING DATE: 2000-04-07
: PRIOR APPLICATION NUMBER: JP 00/280988
: PRIOR FILING DATE: 2000-08-03
: NUMBER OF SEQ ID NOS: 7059
: SOFTWARE: Patentln ver. 3.0
: SEQ ID NO 836
: LENGTH: 1422
: TYPE: DNA
: ORGANISM: Corynebacterium glutamicum
: US-09-738-626-836
```

```

Query Match          8.8%; Score 225.6; DB 9; Length 1422;
Best Local Similarity 56.8%; Pred. No. 1.7e-53;
Matches 477; Conservative 0; Mismatches 354; Indels 9; Gaps 3;

QY 1011 GAAGAGAGCGTGACTGCTGTTACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGT 1070
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 586 GAGGAAACCAACACCGGTGTGACCGGCTGTACCACTTGTGTAAGAGGCGTGTGCT 645
```



```
; TITLE OF INVENTION: METABOLIC PATHWAY PROTEINS
; FILE REFERENCE: BGI-121CP2
; CURRENT APPLICATION NUMBER: US/09/746,660A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 09/606740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 09/603124
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142101
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/148613
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/187970
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: Patent In Vers. 2.0
; SEQ ID NO 101
; LENGTH: 1396
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(1396)
; OTHER INFORMATION: FRXA01371
US-09-746-660A-101

Query Match          7.3%; Score 186; DB 9; Length 1396;
Best Local Similarity 57.3%; Pred. No. 3.2e-42;
Matches 377; Conservative 0; Mismatches 275; Indels 6; Gaps 2;

QY 1011 GAAGAGCGTACTGTTGTTACAGCGCTGTATCAGCTCTCCAAAGCTGGAGCTCTGT 1070
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 698 GAGGAACCAACCACCGGTGACCGCTGTACCACTTCTGTAAGAGCGCTGCTGCT 757

QY 1071 GTCCGCCCATGACGTCATGATGTTTACCAACAGAGTTGATACTTGTACTGC 1130
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 758 TTCCACGATGACGTCACGACGCTGTACCAAGTCCAAAGTTGATAACAAGTACGC 817

QY 1131 TGCCGAATCCATTGATGCGCTGAAGAGACACAGATGATGTTGTTGGGAA 1190
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 818 ACCGCCACTCCCTGATCGACGGCATCAACCGCCACTGACATGCTCATGGCGGCAAG 877

QY 1191 CAAGTGTGTGTGTGCTATGTTAGGTAGGCAAGGCGCTGCTGCTCTCAAGCT 1250
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 878 AACGTGCTGTGCTGCGGTACGGCGATGTCGGAAGGCGCTGCGCTGAGCGCTTCGACGGC 937

QY 1251 CTGGACCAATTGCTACATTACCGAAATGCAACCCCATCTGCTCTGACGGCTGCATG 1310
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 938 CAGGCGCTCGCGTCAAGGTCAAGCTGACCCCAATCAACGCTCTTCAGGCTCTGATG 997

QY 1311 GATGGTTCAGGGTGTAAAGCTAAATGAAGTCAATCCGCAAGTGCATGCTAATACT 1370
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 998 GATGCTACTCTGTGTACCGTGTAGAGCCATGAGAGCGCGCATCATCTGATCACC 1057

QY 1371 TGCACAGAAATAGAATAGTAGTACACGGAGCACTTGATCGCATGAAAAACAGTTGT 1430
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1058 GCGACCGCAACAAGACATCATTTCTTCGACAGATGCTCAAGATGAAGATCAGCT 1117

QY 1431 ATGTAATGCAATATGGCCACTCCACACAGAAATCGATGTCAGCAGCTC--CGACT 1487
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1118 CTGCTGGGCAACATCGGTACCTTGAATGAGATCGATATGCAATCCCTGTTGACCGC 1177

QY 1488 CCGAGCTGACGTGGAGCGAGTACGTTCTCAGGTGACCATGTCATCTGGCCAGATGGC 1547
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1178 GACGACGTACCCGACACGATCAAGCCACAGGTGACGAGTTCACCTTCTCCACCGGT 1237

QY 1548 AAAGAGTTGCTCTCTGCGACAGGCTGCTACTCAATTTGAGCTGCTCCACAGTTCCC 1607
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1238 CGTTCATCATCTGCTCTGTCGGAAGGTGCGCTGTGAACCTTGGCAACGCCACCGGACAC 1297
```

```
QY 1608 ACC---TTTCTTCTGTCATCACAGCCACACAGCGCTTGGCAGCTGATAGACTCT 1662
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1298 CCATCATTTGTCATGTCCAACTCTTTCGCCGATCAGACCATTTGCCGACATGCACTGT 1355

RESULT 15
US-09-754-853A-4/c
; Sequence 4, Application US/09754853A
; Publication No. US20030005491A1
; GENERAL INFORMATION:
; APPLICANT: Hauge, Brian M.
; APPLICANT: Parneil, Laurence D.
; APPLICANT: Parsons, Jeremy D.
; APPLICANT: Wang, Ming Li
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-10(15810)B
; CURRENT APPLICATION NUMBER: US/09/754,853A
; CURRENT FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 60/1174,880
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 1119
; SEQ ID NO 4
; LENGTH: 513509
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (111805)..(113968), (114684)..(115204)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: 318013_region_A3
US-09-754-853A-4

Query Match          6.9%; Score 176.2; DB 9; Length 513509;
Best Local Similarity 55.4%; Pred. No. 9.6e-38;
Matches 428; Conservative 0; Mismatches 333; Indels 12; Gaps 4;

QY 1091 TGATCTGTACCAACAAGAGTTGATTAATCTGTACTGCTGCCGAGAAATCAATTTGA 1150
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 178978 TTAATCTGTGTTTGTGAACAGTTTGACAACCTGTATGGGTGCCGTACCTCTCCCTGA 178919

QY 1151 TGGCTGAAGAGACCAAGATGTGATGTTGGGGAACAAGTGGTGTGTGCTA 1210
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 178918 TGGTCTCATGAGGCTACCGATGTATGATGTGGAAGAGTGGCTGTGTGCTGATA 178859

QY 1211 TGTGAGTAGGCAAGGCTGTGCTGTCTCAAAAGCTCTTGAGCAATGCTACAT 1270
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 178858 TGTGATGTGGAAGGTTGTGCTGCTGCAATGAAGCAAGCTGTGCTGCTCATCGT 178799

QY 1271 TACCGAATGCAACCCCATCTGTGCTGACAGGCTGCATGATGGGTTCAGGGTGTAA 1330
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 178798 GACCGAGATGTATCCCATCTGTGCTGCAAGCTCATGGAAGGCTTCAGGTTCTGAC 178739

QY 1331 GCTAATGAAGTCAATCCGGCAAGTGTGCTTAATTAATCTGCACAGGAATAGAATGT 1390
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 178738 CTGGAGGATGTGTTCTGTGAGGCTGATATCTTTGTCAACCAACCGGTAAAGACAT 178679

QY 1391 AGTGACACGGGACACTGTGATCGCATGAATAACAGTTGATGTAAGCAATATGGCCA 1450
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 178678 CATCATGTTGACACATGAGAAATGAAGAACAAATGCCATTGTTGCAACATGTGCTCA 178619

QY 1451 CTCCAACACAGAAATGATGTGACACAGCTCCGCACT--CCGAGCTGAGGTGGAGCG 1507
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 178618 CTTGACAAATGAGATGCAATGCTTGGGCTGAGAACTAACCCGGCGTGAAGCGCATCAC 178559

QY 1508 AGTACGTTCTCAGGTGACCATGTATCTGCGCAGATGGCAACGAGTTTC--CTCCT 1564
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 178558 CATCAAGCCCCAAACTGACAGATGGGTCTTCCCTGAGACCAACACCGGATATCTTCT 178499
```

[illegible]

```
Search completed: April 20, 2003, 19:29:31
Job time : 1363 secs
```


QY	858	GGCGAGTCAGAAAGATGACTTCTGGTGTGATTTGACCGCTGTGTGAACATGATGGGTGG	917
Db	896	GGAGAATCAGAAGATGACTTTTGGTGTGCATTGACAGATGCGTGAATGTGGAGGGCTGG	955
QY	918	CAGGCCAACATGATCCTGGATGATGGGGAGACTTAAACCACTGGGTTTATAAGAGTAT	977
Db	956	CAGCCAAACATGATTAATGATGATGAGAGATCTGACTCACTGATTTATAAAAAGTAT	1015
QY	978	CCAAACGTGTTTAAGAGATCCGAGGCATTGTGGAAGAGAGCGTGACTGTTCTACAGG	1037
Db	1016	CCCAACATGTTTAAAAAATCAAAAGGCATAGTCGAGGAGAGTGTACTGGAGTCCATAGG	1075
QY	1038	CTGTATCAGCTCTCCAAGCTGGGAAGCTCTGTGTCCGGCCATGAACGTCAATGATCT	1097
Db	1076	CTGTACCAACTGTCCAAGCTGGGAAGCTGTGTGTCCCGCCATGAACGTCAATGACTCA	1135
QY	1098	GTTACCAACAGAAAGTTGATTACTGTACTGCTGCGGAGAATCCATTTTGGATGGCCTG	1157
Db	1136	GTCACCTAAGCAGAAATTTGATTAACCTCTACTGTTGCCGTGAATCTATTCTTGACGGACTT	1195
QY	1158	AAGAGGACCACAGATGTGATGTTTGGTGGGAACAAGTGGTGTGTGGCTATGTTAG	1217
Db	1196	AAAAGACACAGATATGATGTTTGGTGGAAAGCAGTGGTGTCTGTGCTATGAGAG	1255
QY	1218	GTAAGCAAGGCGTGTGCTGCTGCTCTCTCAAGCTCTTGAGCAATTGTCTACATTACGAA	1277
Db	1256	GTCGGAAGGGGTGTGGCTGTCTCTGAAGCCATGGGCTCCATTGTGTACGTACTGAG	1315
QY	1278	ATCGACCCCATCTGTGCTCTGACAGGCCCTGCATGATGGGTTACAGGGTGTAAAGCTAAT	1337
Db	1316	ATTGACCCCATCTGTGCTCTGCAAGCCTGTATGGATGGATTCGGACTAGTGAAGCTGAT	1375
QY	1338	GAAATCATCCGGCAAGTCGATGTCGTAATACTTGACAGGAATAAGAAATGATGACA	1397
Db	1376	GAAATCATCCGACAGGTGACATTTGTTATTACTTGACAGGGGAACAAGAAATGTGTAAC	1435
QY	1398	CGGAGCAGCTTGGATGCGATGAAAAACAGTTGTATGATGCAATATGGGCCACTCCAC	1457
Db	1436	AGAGAGCAGCTTGGACCGTATGAAGAATAGCTGATCGTTGTAAACATGGGCCATTCAC	1495
QY	1458	ACAGAAATCGATGTGACACAGCCTCCGCACTCCGAGCTGACGTGGAGCGAGTAGC- TTC	1516
Db	1496	ACCGAAATGATGTGGCCAGCCTGCGGACACCAAACTGACCTGGAGAGAGTAAGATTC	1555
QY	1517	TCAGGTGACCATGTCTCTGGCCAGAT- GGCAACGAGTTGCTCCTCGCCAGA- GGGT	1574
Db	1556	CCAATTTGATCATGTGATATGCGCTGATGGGCAAGAGATAGTATTGCTGGCAGAGGGGC	1615
QY	1575	CGTCTACTCAATTTGAGCTGCTCCACAGTTCACACTTTGTTCTGTCCATCACAGCCACA	1634
Db	1616	CGTTTAATGAATCTCAGGTGCTCCACAGTGCCTPACATTTGTGCTCTCCATCAGCTTACC	1675
QY	1635	ACACAGGCTTTGGCACTGATAGAAGCTATATAATGACCCGAGGGGCGATACAGCAGGAT	1694
Db	1676	ACTCAGGCTCTCGCCTTGATAGAAGCTTTACAATGCTCCTGAGGGTCCGTATAAGCAGAT	1735
QY	1695	GTTGACTGTGCTTCTTAAGAAAAATGGATGAATACGTTGCCAGCTTGATCTGCCATCATTT	1754
Db	1736	GTTGACTGTGCTTACCAGAAAAATGGATGATGATGTGGCAAGCCTTACACCTGCCACTTT	1795
QY	1755	GATGCCCACTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGACTCAACAAAAAT	1814
Db	1796	GATGCCCACTGACAGAGCTTAACAGATGAACAAGCAAGATATCTGGACTCAACAGAAAT	1855
QY	1815	GGGCATTCAAACCTAATATATACAGATACTAATGAGCCATACTACCAAGAGCAGTCCA	1874
Db	1856	GGACCCCTCAAGCCTTAAGCTACTACAGGATTTAAGTTCCTGTAATGCAAGCCAGAGACAT	1915
QY	1875	CCTGAACACACACTCTA 1892	
Db	1916	GTTAAGAAATAGAACCA 1933	

[illegible]

QY 1737 TTGCATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATAT 1796
|||||
Db 380 TTGCATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATAT 439
QY 1797 CTGGGACTCAACAAAAATGGGCCATTCAAAACCTAATTATTACAGATACTAATGGACATA 1856
|||||
Db 440 CTGGGACTCAACAAAAATGGGCCATTCAAAACCTAATTATTACAGATACTAATGGACATA 499
QY 1857 CTACCAAGGACAGCTCCACCTGAAACACACACTAAGAATAATTTTAAAGATACTT 1916
|||||
Db 500 CTACCAAGGACAGCTCCACCTGAAACACACACTAAGAATAATTTTAAAGATACTT 559
QY 1917 TTATTTCTCTTACTCTCTCTCTCTGATTTTCTCTATATTTTCAATCTCTGTTT 1976
|||||
Db 560 TTATTTCTCTTACTCTCTCTCTCTGATTTTCTCTATATTTTCAATCTCTGTTT 619
QY 1977 CATCTCATATCCAGTTCTGACAGACACAGAACTTGCTCATGGCTCTTACATGA 2036
|||||
Db 620 CATCTCATATCCAGTTCTGACAGACACAGAACTTGCTCATGGCTCTTACATGA 679
QY 2037 AATAGAAGTTCAGGCTCCCTCACTCTAGTCACTAAGAAGATTTTACTCCCCAGCCCA 2096
|||||
Db 680 AATAGAAGTTCAGGCTCCCTCACTCTAGTCACTAAGAAGATTTTACTCCCCAGCCCA 739
QY 2097 GAAAGGTGATTTCTCTCTTACCATTTCTGGGACTTGTCTTATTTAGTACCTTAT 2156
|||||
Db 740 GAAAGGTGATTTCTCTTACCATTTCTGGGACTTGTCTTATTTAGTACCTTAT 798
QY 2157 AACAGGAATGCTAAGTACCTTCTGTGAACAATCTGCAATGTCTAATCGCCTTAA 2216
|||||
Db 799 AACAGGAATGCTAAGTACCTTCTGTGAACAATCTGCAATGTCTAATCGCCTTAA 858
QY 2217 AAGAGCCATTTCTTACCTGCTGAATCAGTGTCTTCACTTCTTACAGAGAAGAGGA 2276
|||||
Db 859 AAGAGCCATTTCTTACCTGCTGAATCAGTGTCTTCACTTCTTACAGAGAAGAGGA 918
QY 2277 TGGTACCTACCCGCGAGGTAGTATGTGGTGTG--CATGTAATTTCCCTTAGAA 2334
|||||
Db 919 TGGTACCTACCCGCGAGGTAGTATGTGGTGTG--CATGTAATTTCCCTTAGAA 978
QY 2335 GTTCCAGCCCTGTCTTCTGCGTAAAGGTGTATGTCC--AGTTCAGAGATGTGTATAT 2392
|||||
Db 979 GTTCCAGCCCTGTCTTCTGCGTAAAGGTGTATGTCCAGTTTAAAGAGGGGTTTAT 1038
QY 2393 GAGCATGGCTGT 2405
|||||
Db 1039 GAGCATGGCTGT 1051

RESULT 4
AL527928 938 bp mRNA linear EST 13-FEB-2001
LOCUS AL527928 LTI_NFL003_NBC3 Homo sapiens cDNA clone CS0DC027YC19 5
DEFINITION prime, mRNA sequence.
ACCESSION AL527928
VERSION AL527928.1 GI:12791421
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 938)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
FEATURES
source 1. 938
/organism="Homo sapiens"

/db_xref="taxon:9606"
/clone="CS0DC027YC19"
/clone_1db="LTI_NFL003_NBC3"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : filangeli@life.com
http://fulllength.invitrogen.com"

BASE COUNT 241 a 208 c 275 g 212 t 2 others
ORIGIN

Query Match 36.1%; Score 926.2; DB 9; Length 938;
Best Local Similarity 99.7%; Pred. No. 1.3e-157;
Matches 936; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 289 TCGGGGAGAGCTGAAGCAGGCGCCAGAGAGATCGAGAGCGCCGAGAGTACTCTTCATG 348
|||||
Db 1 TCGGGGAGAGCTGAAGCAGGCGCCAGAGAGATCGAGAGCGCCGAGAGTACTCTTCATG 60
QY 349 CCACCGTCAACCAAGCGCGCCAGAGCAATCCAGTTGCTGATGATGACAGAGTTCA 408
|||||
Db 61 CCACCGTCAACCAAGCGCGCCAGAGCAATCCAGTTGCTGATGATGACAGAGTTCA 120
QY 409 CCAAAATCCCAACCAAACTGGCCGAGATCTTTGCTCGCTCGATCTCACAGTCTCCA 468
|||||
Db 121 CCAAAATCCCAACCAAACTGGCCGAGATCTTTGCTCGCTCGATCTCACAGTCTCCA 180
QY 469 CTGACAGCTACAGTTCAAGTCACTACACAGATAGCTGTGATGATGAGTTCTCCCC 528
|||||
Db 181 CTGACAGCTACAGTTCAAGTCACTACACAGATAGCTGTGATGATGAGTTCTCCCC 240
QY 529 GAGAGAGCAGCAAACTCCAGGGCAGCAGCAATTTCTGTGGAAGAATCAACAGC 588
|||||
Db 241 GAGAGAGCAGCAAACTCCAGGGCAGCAGCAATTTCTGTGGAAGAATCAACAGC 300
QY 589 AGGCAGAAATTTGAGCGCGGGAGATGAGATTCAGAGAGCAAGACATGTCTCTGATT 648
|||||
Db 301 AGGCAGAAATTTGAGCGCGGGAGATGAGATTCAGAGAGCAAGACATGTCTCTGATT 360
QY 649 CACTCAGAAACGTGCTCAGGGGGAGAGACCCCTTGCTGTGCTAAATAGTGGCTGTA 708
|||||
Db 361 CACTCAGAAACGTGCTCAGGGGGAGAGACCCCTTGCTGTGCTAAATAGTGGCTGTA 420
QY 709 CACACATCACAGCCAGACAGCGGTGTTGATTGAGACACTCTGTGCCCTGGGGCTCAGT 768
|||||
Db 421 CACACATCACAGCCAGACAGCGGTGTTGATTGAGACACTCTGTGCCCTGGGGCTCAGT 480
QY 769 GCCGCTGCTGCTTGTACATCTACTCAACTCAGAAATGAAGTACCTGACACTGGCTG 828
|||||
Db 481 GCCGCTGCTGCTTGTWACATCTACTCAACTCAGAAATGAAGTACCTGACACTGGCTG 540
QY 829 AGCCTGAGATTGCAAGTGTGCGCTGGAAGGGCGAGTCAGAGATGACTTCTGCTGTA 888
|||||
Db 541 AGCCTGAGATTGCAAGTGTGCGCTGGAAGGGCGAGTCAGAGATGACTTCTGCTGTA 600
QY 889 TTGACCGCTGTGTGAACATGATGGTGGCAGGCCAACATGATCTCGATGATGGGGAG 948
|||||
Db 601 TTGACCGCTGTGTGAACATGATGGTGGCAGGCCAACATGATCTCGATGATGGGGAG 660
QY 949 ACTTAACCACTGGGTTTATAGAAGTATCCAAAGTGTGTTAAGAAGATCCGAGGATG 1008
|||||
Db 661 ACTTAACCACTGGGTTTATAGAAGTATCCAAAGTGTGTTAAGAAGATCCGAGGATG 720
QY 1009 TGAAGAGAGCGTGAAGTGTGTTACAGGCTGTATCAGCTCTCCAAAGCTGGAGGCTCT 1068
|||||

Db 721 TGAAGAGAGCGTACTGTTTACAGCGTGTATCAGCTCTCCAAAGCTGGGAAGCTCT 780

QY 1069 GTGTTCCGGCCATGAACGTCAATGATTTCTTACCAACAGAGTTTGAATCTGTACT 1128
|||||
Db 781 GTGTTCCGGCCATGAACGTCAATGATTTCTTACCAACAGAGTTTGAATCTGTACT 840

QY 1129 GCTGCCGAGATCCATTTTGGATGGCCCTGAGAGAGACCACAGATGTGATTTTGGTGGGA 1188
|||||
Db 841 GCTGCCGAGATCCATTTTGGATGGCCCTGAGAGAGACCACAGATGTGATTTTGGTGGG- 899

QY 1189 AACAGTGTGTGTGTGTGCTATGTTGAGGTAGGCCAAG 1227
|||||
Db 900 AAMAAGTGTGTGTGTGCTATGTTGAGGTAGGCCAAG 938

RESULT 5
AL551097 921 bp mRNA linear EST 16-FEB-2001
LOCUS AL551097 LTI_NFL006_PL2 Homo sapiens CDNA clone CS0D1066YM07 5
DEFINITION prime, mRNA sequence.
ACCESSION AL551097
VERSION AL551097.1 GI:12888715
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 921)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length CDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
Source 1. 921
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1066YM07"
/clone_11b="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand CDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded CDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 217 a 231 c 302 g 170 t 1 others

ORIGIN

Query Match 35.0%; Score 897.4; DB 9; Length 921;
Best Local Similarity 99.6%; Pred. No. 2.1e-152;
Matches 909; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 69 GCCTGGGCTGCCGACAGACAGAGCGCTGGCCACAGACCTCAGAGCCGAGCAGCTC 128
|||||
Db 1 GCCTGGGCTGCCGACAGACAGAGCGCTGGCCACAGACCTCAGAGCCGAGCAGCTC 60

QY 129 GACGAGGGGGCCGAGAGAGGTGGGGGATCGCTGTGCGAGGGCGCGCGCGGCGAGGC 188
|||||
Db 61 GACGAGGGGGCCGAGAGAGGTGGGGGATCGCTGTGCGAGGGCGCGCGCGGCGAGGC 120

QY 189 GGGCGGGCCAGAGAGGGGGAAGAGCGGGGGCGGGGCTACGCCCTGGCCGGCGCGGC 248
|||||
Db 121 GGGCGGGCCAGAGAGGGGGAAGAGCGGGGGCGGGGCTACGCCCTGGCCGGCGCGGC 180

QY 249 GGGGGAATGTGATGCTGACCGCATGCGCTGCCCGGGGTGCGGGAGGAGCTGAAGCAG 308
|||||

Db 181 CGGGGAATGTGATGCTCTGACGCGATGCCGCTGCCCGGGGTGCGGGAGGAGCTGAAGCAG 240

QY 309 GCCAAGAGATCGAGAGCGCCGAGAGTACTCTCTCATGCGCCACCGTCACCAAGGCCCC 368
|||||
Db 241 GCCAAGAGATCGAGAGCGCCGAGAGTACTCTCTCATGCGCCACCGTCACCAAGGCCCC 300

QY 369 AAGAAGCAATCCAGTTGCTGATGACATGACAGGAGTTCAACAAATTCGCCCAAACT 428
|||||
Db 301 AAGAAGCAATCCAGTTGCTGATGACATGACAGGAGTTCAACAAATTCGCCCAAACT 360

QY 429 GGGCGAAGATCTTGTCTGCTGCTGATGATGAGTTTCTCCCGAGAGAGCAAGCAAC 488
|||||
Db 361 GGGCGAAGATCTTGTCTGCTGCTGATGATGAGTTTCTCCCGAGAGAGCAAGCAAC 420

QY 489 GCATCCTACACAGATAGCTCTGATGATGAGTTTCTCCCGAGAGAGCAAGCAAC 548
|||||
Db 421 GCATCCTACACAGATAGCTCTGATGATGAGTTTCTCCCGAGAGAGCAAGCAAC 480

QY 549 TCCAGGGCAGCAGCAATTTCTGTGTGAAGACATCAAGCAGCAGATTTGGAGCGCCG 608
|||||
Db 481 TCCAGGGCAGCAGCAATTTCTGTGTGAAGACATCAAGCAGCAGATTTGGAGCGCCG 540

QY 609 GAGATTGAGATTGCAGAGCAGACATGTCTGCTCTGATTTCACTCAGGAACGTGCTCAG 668
|||||
Db 541 GAGATTGAGATTGCAGAGCAGACATGTCTGCTCTGATTTCACTCAGGAACGTGCTCAG 600

QY 669 GGGGAGAGCCCTTGCTGCTGCTGCTAAATAGTGGGCTGTACACACATCAGCCACAGCA 728
|||||
Db 601 GGGGAGAGCCCTTGCTGCTGCTGCTAAATAGTGGGCTGTACACACATCAGCCACAGCA 660

QY 729 GCGGTGTGATTGAGACACTCTGTCCTGGGGCTGAGTCCGCTGCTGCTGTAAC 788
|||||
Db 661 GCGGTGTGATTGAGACACTCTGTCCTGGGGCTGAGTCCGCTGCTGCTGTAAC 720

QY 789 ATCTACTCACTCAGATGAAGTAGCTGACGACACTGCTGAGGCTGAGTTGCACTTC 848
|||||
Db 721 ATCTACTCACTCAGATGAAGTAGCTGACGACACTGCTGAGGCTGAGTTGCACTTC 780

QY 849 GCTTGAAGGGCGAGTCAAGATGACTTCTGCTGCTGATTTGACCCGCTGTGTAACATG 908
|||||
Db 781 GCTTGAAGGGCGAGTCAAGATGACTTCTGCTGCTGATTTGACCCGCTGTGTAACATG 840

QY 909 GATGGTGGCAGGCCAATGATGCTGATGATGAGGGGAGACTTAACCACTGGTTAT 968
|||||
Db 841 GATGGTGGCAGGCCAATGATGCTGATGATGAGGGGAGACTTAA-CCACTGGTTAT 899

QY 969 AAGAAGTATCAA 981
|||||
Db 900 AARAAGTATCAA 912

RESULT 6
AL520703/c 957 bp mRNA linear EST 13-FEB-2001
LOCUS AL520703 LTI_NFL004_NBC2 Homo sapiens CDNA clone CS0DB002YL22 3
DEFINITION prime, mRNA sequence.
ACCESSION AL520703
VERSION AL520703.1 GI:12784196
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 957)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length CDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
Source 1. 957
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DB002YL22"
/clone_1lb="LTI_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : filiang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 299 a 176 c 227 g 254 t 1 others
ORIGIN

Query Match 35.0%; Score 896.2; DB 9; Length 957;
Best Local Similarity 99.1%; Pred. No. 3.4e-152;
Matches 942; Conservative 1; Mismatches 4; Indels 4; Gaps 4;

QY 1572 GGTGCTCTACTCAATTTGAGCTGCTCCACAGTTCACCTTTGTTCTGTCATCACAGCC 1631
|||||
Db 947 GGTGCTCTACTCAATTTGAGCTGCTCCACAG-TCCACACCTTTGTTCTGTCATCACAGCC 889
QY 1632 ACAACACAGGCTTTGGCAGTGAATCTATATATGACCCGAGGGGCGATACAAGCAG 1691
|||||
Db 888 ACAACACAGGCTTTGGCAGTGAATCTATATATGACCCGAGGGGCGATACAAGCAG 829
QY 1692 GATGTGTACTGCTCTCCTAAGAAATGATGAATAGCTGCGCAGCTGTCATCTGCCATCA 1751
|||||
Db 828 GATGTGTACTGCTCTCCTAAGAAATGATGAATAGCTGCGCAGCTGTCATCTGCCATCA 769
QY 1752 TTGTATGCCCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAACAAA 1811
|||||
Db 768 TTGTATGCCCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAACAAA 709
QY 1812 AATGGGCAATTCAAACCTAATTTATACAGATCTATATGAGCATCTACCAAGGACAGT 1871
|||||
Db 708 AATGGGCAATTCAAACCTAATTTATACAGATCTATATGAGCATCTACCAAGGACAGT 649
QY 1872 CCACCTGAACACACACACTCTAAGAAATATTTTAAAGATTAACCTTTTCTCTTCTTAC 1931
|||||
Db 648 CCACCTGAACACACACACTCTAAGAAATATTTTAAAGATTAACCTTTTCTCTTCTTAC 589
QY 1932 TCCTTTCCTCTGTGATTTTTCCTAATAATTCATCTCTGTTTTCATCTCATATATCCAA 1991
|||||
Db 588 TCCTTTCCTCTGTGATTTTTCCTAATAATTCATCTCTGTTTTCATCTCATATATCCAA 529
QY 1992 GTTCTGACAGACACACAGGACTGCTTCATGCGCTCTTAGATGAATAGAAAGTTCAAGG 2051
|||||
Db 528 GTTCTGACAGACACACAGGACTGCTTCATGCGCTCTTAGATGAATAGAAAGTTCAAGG 469
QY 2052 TCCCTCACTCTAGTCACTAAGAAAGATTTTACTCCCCAGCCGAGAAAGTGATCTTC 2111
|||||
Db 468 TCCCTCACTCTAGTCACTAAGAAAGATTTTACTCCCCAGCCGAGAAAGTGATCTTC 410
QY 2112 TCTTTACATTTCTGGGACTTTAGTCTTAATTAAGTACCTTATTAACAGGAAATGCTAA 2171
|||||
Db 409 TCTTTACATTTCTGGGACTTTAGTCTTAATTAAGTACCTTATTAACAGGAAATGCTAA 350
QY 2172 GGTACCTTCTCTGTGGAACAATCTGCAATGTCTAATATCGCCTTAAAGAGCCATTTCTT 2231
|||||
Db 349 GGTACCTTCTCTGTGGAACAATCTGCAATGTCTAATATCGCCTTAAAGAGCCATTTCTT 290
QY 2232 AGCTGCTGAATATCAGTCTCTTTCATCTCTTACAGAGAAAGGAGGATGTTACTACCCGGC 2291
|||||
Db 289 AGCTGCTGAATATCAGTCTCTTTCATCTCTTACAGAGAAAGGAGGATGTTACTACCCGGC 230
QY 2292 AGGTAGGTTAGATGTGGGTGTCATGTTAATTTCCCTTAGAAGTTCAGAGCCCTGTTTC 2351

|||||
Db 229 AGGTAGGTTAGATGTGGGTGTCATGTTAATTTCCCTTAGAAGTTCAGAGCCCTGTTTC 170
QY 2352 CTGCGTAAAGGTGTATGTCCAGTTCAGAGATGTATATATGAGCATGGCTTGTAAAGT 2411
|||||
Db 169 CTGCGTAAAGGTGTATGTCCAGTTCAGAGATGTATATATGAGCATGGCTTGTAAAGT 110
QY 2412 CAGAGAGCCCACTTGATTTATAGTATAGCCCTTCTCCACTCCACCCAGACTTGCTCAT 2471
|||||
Db 109 CAGAGAGCCCACTTGATTTATAGTATAGCCCTTCTCCACCCACCCAGACTTGCTCAT 50
QY 2472 TTTTCGAGTTTAACTAGACTACACTCTATTTGAGTTAATTTGTCTC 2522
|||||
Db 49 TTTTCGAG-TTTTAACTAGACTACACTCTA-TTGAGTTTAAATTTGTCTC 1

RESULT 7
AL527927/c 950 bp mRNA linear EST 13-FEB-2001
LOCUS AL527927 LTI_NFL003_NBC3 Homo sapiens cDNA clone CS0DC027YC19 3
DEFINITION prime, mRNA sequence.
ACCESSION AL527927
VERSION AL527927.1 GI:12791420
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 950)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..950
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DC027YC19"
/clone_1lb="LTI_NFL003_NBC3"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : filiang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 291 a 177 c 223 g 255 t 4 others
ORIGIN

Query Match 34.5%; Score 884.4; DB 9; Length 950;
Best Local Similarity 99.1%; Pred. No. 4.6e-150;
Matches 907; Conservative 2; Mismatches 4; Indels 2; Gaps 2;

QY 1572 GGTGCTCTACTCAATTTGAGCTGCTCCACAGTTCACCTTTGTTCTGTCATCACAGCC 1631
|||||
Db 913 GGTGCTCTACTCAATTTGAGC-GCTCCACAGTCCACCTTTGTTCTGTCATCACAGCC 855
QY 1632 ACAACACAGGCTTTGGCAGTGAATCTATATATGACCCGAGGGGCGATACAAGCAG 1691
|||||
Db 854 ACAACACAGGCTTTGGCAGTGAATCTATATATGACCCGAGGGGCGATACAAGCAG 795
QY 1692 GATGTGTACTGCTCTCCTAAGAAATGATGAATAGCTGCGCAGCTTGCATCTGCCATCA 1751
|||||
Db 794 GATGTGAACCTGCTCTCCTAAGAAATGATGAATAGCTTGCATCTGCCATCA 735

```

QY 1752 TTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATCTGGACTCAACAA 1811
    |||||||
Db 734 TTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATCTGGACTCAACAA 675
QY 1812 AATGGCCATTCAACCTAATTAATACAGATACTAATGACCACTACTACCAAGACAGCAGT 1871
    |||||||
Db 674 AATGGCCATTCAACCTAATTAATACAGATACTAATGACCACTACTACCAAGACAGCAGT 615
QY 1872 CCACTGAACACACACACTCTAAGAAATATTTTAAGATACTTTTCTCTCTAC 1931
    |||||||
Db 614 CCACTGAACACACACACTCTAAGAAATATTTTAAGATACTTTTCTCTCTAC 555
QY 1932 TCCCTTCTCTGATTTTCTCTAATTTCAATCTCTGTTTTCATCTCATTCACAA 1991
    |||||||
Db 554 TCCCTTCTCTGATTTTCTCTAATTTCAATCTCTGTTTTCATCTCATTCACAA 495
QY 1992 GTTTCGACAGACACACAGAACTGCTTCATGCTCTTTAGATGAATAGAACTCAGG 2051
    |||||||
Db 494 GTTTCGACAGACACACAGAACTGCTTCATGCTCTTTAGATGAATAGAACTCAGG 435
QY 2052 TCCCTCACTCTAGTCACTAAGAAAGATTTTACTCCCCAGCCCAAGAGGTGATCTTC 2111
    |||||||
Db 434 TTCTCACTCTAGTCACTAAGAAAGATTTTACTCTCCAGCCCAAGAGGTGATCTTC 376
QY 2112 TCTTACCATTTCTGGGACTTGTCTTAATTAAGTACCTTATTAACAGAAATGCTAA 2171
    |||||||
Db 375 TCTTACCATTTCTGGGACTTGTCTTAATTAAGTACCTTATTAACAGAAATGCTAA 316
QY 2172 GGTACCTTCTCTGTGAACAATCTGAATGTCTAATCGCCTTAAAGAGCCATTTCTT 2231
    |||||||
Db 315 GGTACCTTCTCTGTGAACAATCTGAATGTCTAATCGCCTTAAAGAGCCATTTCTT 256
QY 2232 AGCTGCTGAATCAGTGTCTTCTCTCTTCAGAGAGCAGGAGTGTACCTACCCGGC 2291
    |||||||
Db 255 AGCTGCTGAATCAGTGTCTTCTCTCTTCAGAGAGCAGGAGTGTACCTACCCGGC 196
QY 2292 AGTAGGTAGATGTGGGTGCTGATTAATTTCCCTTAGAAGTCCAGCCCTGTTTC 2351
    |||||||
Db 195 AGTAGGTAGATGTGGGTGCTGATTAATTTCCCTTAGAAGTCCAGCCCTGTTTC 136
QY 2352 CTGCGTAAAGGTGTATGTCCAGTTCAGAGATGTGTATATAGAGCATGCTGTTAGAT 2411
    |||||||
Db 135 CTGCGTAAAGGTGTATGTCCAGTTCAGAGATGTGTATATAGAGCATGCTGTTAGAT 76
QY 2412 CAGAGGCCCACTTGATTTATAGTATAGCCCTTCTCCACTCCACACAGACTGCTCAT 2471
    |||||||
Db 75 CAGAGGCCCACTTGATTTATAGTATAGCCCTTCTCCACTCCACACAGACTNGCTCAT 16
QY 2472 TTTTCGAGTTTAA 2486
    |||||||
Db 15 TTTTCGAGTTTAA 1

```

```

RESULT 8
LOCUS AL520704 891 bp mRNA linear EST 13-FEB-2001
DEFINITION AL520704 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB002YL22 5
prime, mRNA sequence.
ACCESSION AL520704
VERSION AL520704.1 GI:12784197
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 891)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France

```

```

FEATURES             Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
    source            location/Qualifiers
                        1. 891
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="CS0DB002YL22"
                        /clone_1lb="LTI_NFL004_NBC2"
                        /sex="male"
                        /tissue_type="neuroblastoma cells"
                        /lab_host="DH10B"
                        /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen.9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT          194 a 238 c 299 g 158 t 2 others
ORIGIN
Query Match          34.1%; Score 873; DB 9; Length 891;
Best Local Similarity 99.5%; Pred. No. 5.4e-148;
Matches 873; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 GGGCGGGCAGGTGCGAGCTCGGAGCTGCTGCTGCTCTCTTGTGGCCACCGTCGCT 60
    |||||||
Db 15 GGGCGGGCAGGTGCGAGCTCGGAGCTGCTGCTGCTCTCTTGTGGCCACCGTCGCT 74
QY 61 GTCCGGCTGCTTGGGCTGCCGAACAGACAGAGCGGTGGCCACAGCACCCTCAGAACCCGA 120
    |||||||
Db 75 GTCCGGCTGCTTGGGCTGCCGAACAGACAGAGCGGTGGCCACAGCACCCTCAGAACCCGA 134
QY 121 CGCAGCTCGACGAGGGGGCGGCGAGAGGGTGGCGATCGCGTGTGGAGGGCGCGCGC 180
    |||||||
Db 135 CGCAGCTCGACGAGGGGGCGGCGAGAGGGTGGCGATCGCGTGTGGAGGGCGCGCGC 194
QY 181 GGGCAGGGCGGGCGGCCAGAGGGGGAAGAGGGGGGGCGGGGTGACCGCTGGCC 240
    |||||||
Db 195 GGGCAGGGCGGGCGGCCAGARGGGGAARAGGGGGGGCGGGGTGACCGCTGGCC 254
QY 241 GGGCGGGCGGGGATGTGATGCTGACCGGATGCCCTGCCCCGGGTGGGGAGAGC 300
    |||||||
Db 255 GGGCGGGCGGGGATGTGATGCTGACCGGATGCCCTGCCCCGGGTGGGGAGAGC 314
QY 301 TGAAGCAGGCCAAGAGATCGAGAGCGCCGAGAAGTACTCCTCATGCGCACCCCTCACCA 360
    |||||||
Db 315 TGAAGCAGGCCAAGAGATCGAGAGCGCCGAGAAGTACTCCTCATGCGCACCCCTCACCA 374
QY 361 AGCGGCCCAAGAACCAATCCAGTTGCTGATGATGACATGACAGAGTTCACCAATTCCTCCA 420
    |||||||
Db 375 AGCGGCCCAAGAACCAATCCAGTTGCTGATGATGACATGACAGAGTTCACCAATTCCTCCA 434
QY 421 CCAAACTGGCCGAGATCTTGTCTCGCTGATCTCACAGTCTCCACTGACAGCTACA 480
    |||||||
Db 435 CCAAACTGGCCGAGATCTTGTCTCGCTGATCTCACAGTCTCCACTGACAGCTACA 494
QY 481 GTTCAGCTGATCTTACACAGATAGCTGTGATGATGAGGTTTCTCCCGAGAGAGCAGC 540
    |||||||
Db 495 GTTCAGCTGATCTTACACAGATAGCTGTGATGATGAGGTTTCTCCCGAGAGAGCAGC 554
QY 541 AAACCAACTCGAAGGCGACAGCAATTTCTGTGGAAGAATCAAGCAGGAGAAATTG 600
    |||||||
Db 555 AAACCAACTCGAAGGCGACAGCAATTTCTGTGGAAGAATCAAGCAGGAGAAATTG 614
QY 601 GACGCCGGAGATGAGATTCAGAGACAGACATGTCTGCTGATTTCTACTCAGGAAC 660
    |||||||
Db 615 GACGCCGGAGATGAGATTCAGAGACAGACATGTCTGCTGATTTCTACTCAGGAAC 674
QY 661 GTGCTCAGGGGAGAGCCCTTGCTGTCTAAATAGTGGGCTGTACACATCACAG 720
    |||||||

```

Db 675 GTGCTCAGGGGGAGAGCCCTTGCTGCTAAATAGTGGGCTGTACACACATCACAG 734

QY 721 CCCAGACAGCGGTGTGATGTAGACACTCTGTGCCCTGGGGGCTAGTCCGCTGTCTG 780

Db 735 CCCAGACAGCGGTGTGATGTAGACACTCTGTGCCCTGGGGGCTAGTCCGCTGTCTG 794

QY 781 CTTGTAACTCTACTCAACTCAGATGAAGTAGCTGCAGCAGTGGCTGAGGCTGAGTTG 840

Db 795 CTTGTAACTCTACTCAACTCAGATGAAGTAGCTGCAGCAGTGGCTGAGGCTGAGTTG 854

QY 841 CAGTGTGCTTGAAGGGGAGTCAAGATGACTT 877

Db 855 CAGTGTGCTTGAAGGGGAGTCAAGATGACTT 891

RESULT 9

AL577470/c 917 bp mRNA linear EST 16-FEB-2001

LOCUS AL577470 LTI_NFL006_PL2 Homo sapiens cDNA clone CSODI087YL18 3

DEFINITION prime, mRNA sequence.

ACCESSION AL577470

VERSION AL577470.1 GI:12940631

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 917)

AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source

1. 917

location/Qualifiers

organism="Homo sapiens"

db_xref="taxon:9606"

clone="CSODI087YL18"

clone_1lb="LTI_NFL006_PL2"

tissue_type="placenta"

/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371

Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 280 a 171 c 210 g 247 t 9 others

ORIGIN

Query Match 33.9%; Score 868.6; DB 9; Length 917;

Best Local Similarity 97.3%; Pred. No. 3.4e-147;

Matches 885; Conservative 7; Mismatches 17; Indels 1; Gaps 1;

QY 1594 GCTCCACAGTCCACCTTGTGTCATCACAGCCACACACAGGCTTTGGCACTGA 1653

Db 917 GCTCCACAGTCCACCTTGTGTCATCACAGCCACACACAGGCTTTGGCACTGA 858

QY 1654 TAGAAGCTATATGACCCGAGGGGCGATACAGCAGAGATGTACTTGCCTCTAAGA 1713

Db 857 TAGAAGCTATATGACCCGAGGGGCGATACAGCAGAGATGTACTTGCCTCTAAGA 798

QY 1714 AATGATGAATACGTGCGACGCTGCATCGCCATCATTTGATGCCACCTTACAGAGC 1773

Db 797 AATGATGAATACGTGCGCGCTGCTGCTGCATCATTTGATGCCACCTTACAGAGC 738

QY 1774 TGACAGATGACCAAGCAAAATATCTGGAGCTCAACAAAATGGCCATTCAAACTAAT 1833

Db 737 TSACAGATGACCAAGCAAAATATCTGGGACTCAACAAAATGGCCATTCAAACTAAT 678

QY 1834 ATTACAGATACTAATGAGCACTACTACCAAGGACAGTCCACCTGAAACCACACTTAA 1893

Db 677 ATTACAGATACTAATGAGCACTACTACCAAGGACAGTCCACCTGAAACCACACTTAA 618

QY 1894 AGAATATTTTAAAGTAACCTTTATTTCTTCTTACTCTCTTCTCTGATTTTTC 1953

Db 617 AGAATATTTTAAAGTAACCTTTATTTCTTCTTACTCTCTTCTCTGATTTTTC 558

QY 1954 CTATAATTTCAATCTCTGTTTTTTCATCTCATATATCAAGTTCTGCAGACCACAGAAAC 2013

Db 557 CTATAATTTCAATCTCTGTTTTTTCATCTCATATATCAAGTTCTGCAGACCACAGAGGAC 498

QY 2014 TTGCTTCATGGCTCTTACATGAATAGAGTTCAGGTCCTCCTACTAGTCACTTAAAG 2073

Db 497 TTGCTTCATGGCTCTTACATGAATAGAGTTCAGGTCCTCCTACTAGTCACTTAAAG 438

QY 2074 AAGGATTTACTCCCCCAGCCAGAAAGGTGATTTCTTCTTACCATTTCTGGGACTT 2133

Db 437 AAGGATTTNACTCTCCAGCCAGAAAGGTGATTTCTTCTTACCATTTCTGGGACTT 379

QY 2134 TAGCTTAATTAGGTACCTTATTAACAGGAAATGCTAAGTACCTTCTGTGGAACAAT 2193

Db 378 TAGCTTAATTAGGTACCTTATTAACAGGAAATGCTAAGTACCTTCTGTGGAACAAT 319

QY 2194 CTGCAATGTCTAAATCGCCTTAAAGAGCCCATTTCTTACTGCTGAATCAGTGTCTT 2253

Db 318 CTGCAATGTCTAAATCGCCTTAAAGAGCCCATTTCTTACTGCTGCTGAATCAGTGTCTT 259

QY 2254 TCACTTCTCAGAGAGAGGAGTGTACCTACCCGGCAGTAGGTAGATGTGGTGTGT 2313

Db 258 TCACTTCTCAGAGAGAGGAGTGTACCTACCCGGCAGTAGGTAGATGTGGTGTGT 199

QY 2314 GCATGTTAATTTCCCTTAGAGTCCCAAGCCCTGTTCCTGCGTAAAGGTGTATGTCCA 2373

Db 198 GCATGTTAATTTCCCTTAGAGTCCCAAGCCCTGTTCCTGCGTAAAGGTGTATGTCCA 139

QY 2374 GTTCAGAGATGTGTATATAGCATGGCTGTGTTAAGATCAGAGGCCCACTTGGATTAT 2433

Db 138 GTTCAGAGATGTGTATATAGCATGGCTGTGTTAAGATCAGAGGCCCACTTGGATTAT 79

QY 2434 AGTATAGCCCTTCTCCACTCCACCCAGACTTGTGTCATTTTTCAGATTTTACTAGACT 2493

Db 78 AGTATAGCCCTTCTCCACTCCACCCAGACTTGTGTCATTTTTCAGATTTTACTAGACT 19

QY 2494 ACACCTCTATT 2503

Db 18 ACAGTCTATT 9

RESULT 10

AL574762/c 916 bp mRNA linear EST 16-FEB-2001

LOCUS AL574762 LTI_NFL006_PL2 Homo sapiens cDNA clone CSODI066YM07 3

DEFINITION prime, mRNA sequence.

ACCESSION AL574762

VERSION AL574762.1 GI:12935277

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 916)

AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source

1. 916

|||||
Db 181 ACATCACAGCCACAGACGCGGTGATTGAGACACTCTGTGCGCTGGGGCTCAGTGCC 240
QY 772 GCTGGTCTGCTGTGAACATCTACTCACTCAGAAATGAAGTAGCTGCAGCACTGGCTGAGG 831
|||||
Db 241 GCTGGTCTGCTGTGAACATCTACTCACTCAGAAATGAAGTAGCTGCAGCACTGGCTGAGG 300
QY 832 CTGGAGTTGACAGTGTTCGCTTGAAGGGCGAGTCAAGAAATGACTTCTGCTGATATG 891
|||||
Db 301 CTGGAGTTGACAGTGTTCGCTTGAAGGGCGAGTCAAGAAATGACTTCTGCTGATATG 360
QY 892 ACCGCTGTGTAACATGGATGGGTGGCAGGGCCACATGATCCTGGATGATGGGGAGACT 951
Db 361 ACCGCTGTGTAACATGGATGGGTGGCAGGGCCACATGATCCTGGATGATGGGGAGACT 420
QY 952 TAACCCACTGGGTTTATAGAAGATCCAAACGCTGTTAAGAAGATCCGAGGCACTTGTG 1011
|||||
Db 421 TAACCCACTGGGTTTATAGAAGATCCAAACGCTGTTAAGAAGATCCGAGGCACTTGTG 480
QY 1012 AAGAGAGCGTACTGTTTCAAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTG 1071
|||||
Db 481 AAGAGAGCGTACTGTTTCAAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTG 540
QY 1072 TTCCGGCCATGAACGTCAATGATCTGTACCAACAGAGATTGATACTGTACTGCT 1131
|||||
Db 541 TTCCGGCCATGAACGTCAATGATCTGTACCAACAGAGATTGATACTGTACTGCT 600
QY 1132 GCCGAGAATCCATTTTGGATGGCTGAAGAGACACAGATGTGATGTTGGTGGGAAC 1191
|||||
Db 601 GCCGAGAATCCATTTTGGATGGCTGAAGAGACACAGATGTGATGTTGGTGGGAAC 660
QY 1192 AAGTGGTGGTGTGGCTATGTTGAGGTAGGCAAGGGCTGTGCTGCTCTCAAGCTC 1251
|||||
Db 661 AAGTGGTGGTGTGGCTATGTTGAGGTAGGCAAGGGCTGTGCTGCTCTCAAGCTC 720
QY 1252 TTGGAGCAATTTGTCTACATTTACCGAAATCGACCCCATCTGTCTGTCAGGCTGCATGG 1311
|||||
Db 721 TTGGAGCAATTTGTCTACATTTACCGAAATCGACCCCATCTGTCTGTCAGGCTGCNTGN 780
QY 1312 ATGGG--TTCAAGGTGTAAAGCTAAATGAAGTATCCGCAAGTGTCTGTAATAAC 1369
|||||
Db 781 ATGGGGTTCAAGGNTGTAAGCTAAATGAAGTATCCGCAAGTGTCTGTAATAAC 840
QY 1370 TTGCACACGAAATGAAGATTTAGTGCACACGGGAGCACTT-GGATCGCATGAAAAACAGTT 1428
|||||
Db 841 TTGCACACGAAATGAAGATTTAGTGCACACGGGAGCACTTGGGATCGCATGAAAAACAGTT 900
QY 1429 GTATC-GTATGCAATATGGGCCACTCCACACAGAAA 1464
|||||
Db 901 GTATCGTATGCAATATGGGCCANCTCAACACACANA 937

RESULT 12
AL519571 865 bp mRNA linear EST 13-FEB-2001
LOCUS AL519571 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB004Y104 5
DEFINITION prime, mRNA sequence.
ACCESSION AL519571
VERSION AL519571.1 GI:12783064
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 865)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES Location/Qualifiers

source 1. 865
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DB004Y104"
/clone_1lb="LTI_NFL004_NBC2"
/sex="male"
/issue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dt) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by life technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 193 a 231 c 284 g 155 t 2 others
ORIGIN

Query Match 33.4%; Score 856.2; DB 9; Length 865;
Best Local Similarity 99.2%; Pred. No. 6e-145;
Matches 858; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 8 GCAGGTCGAGCTCGAGCTGCTCTCTGTTCTCTTGTGCGCACCGCTGCTGCTGCGG 67
Db 1 GCAGGTCGAGCTCGAGCTGCTCTCTGTTCTCTTGTGCGCGCGCTGCTGCTGCTGCGG 60
QY 68 TGCCTTGGGCTGCCAAGACAGACAGCGGTGGCCACAGCACCCTCAGAAAGCCGAGCAGCT 127
Db 61 TGCCTTGGGCTGCCAAGACAGACAGCGGTGGCCACAGCACCCTCAGAAAGCCGAGCAGCT 120
QY 128 CGACGAGGGCGCGCGAGAGGGTGGCGATGCGTTCGAGAGGGCGCGCGCGAGG 187
Db 121 CGACGAGGGCGCGCGAGAGGGTGGCGATGCGTTCGAGAGGGCGCGCGCGAGG 180
QY 188 CGGGCGGCGCGAGAGGGGAAAGAGCGGGGGGGGTCAGCCGCTGGCGGGCGG 247
Db 181 CGGGCGGCGCGAGAGGGGAAAGAGCGGGGGGGTCAAGCGCTGGCGGGCGG 240
QY 248 CGGGGAATGTCGATGCTGACCGCGATGCCGCTGCCGGGGTGGGGAGGAGCTGAAGCA 307
Db 241 CGGGGAATGTCGATGCTGACCGCGATGCCGCTGCCGGGGTGGGGAGGAGCTGAAGCA 300
QY 308 GGGCAAGGAGATCGAGAGCGCGAGAGTACTCTTCATGCGCCACCGTCAACCAAGCGGC 367
Db 301 GGGCAAGGAGATCGAGAGCGCGAGAGTACTCTTCATGCGCCACCGTCAACCAAGCGGC 360
QY 368 CAAGAAGCAATTCAGTTTGTCTGATGACATGACAGAGTTCACCAATTCGCCCAAAAC 427
Db 361 CAAGAAGCAATTCAGTTTGTCTGATGACATGACAGAGTTCACCAATTCGCCCAAAAC 420
QY 428 TGGCCGAAGATCTTGTCTCGCTCGATCTCACAGTCTCCACTGACAGCTACAGTTTCAAGC 487
Db 421 TGGCCGAAGATCTTATCTCGCTCGATCTCACAGTCTCCACTGACAGCTACAGTTTCAAGC 480
QY 488 TGCATCTACACAGATAGCTCTGATGATGAGGTTTCTCCCGAGAGAGCAGCAAAACCA 547
Db 481 TGCATCTACACAGATAGCTCTGATGATGAGGTTTCTCCCGAGAGAGCAGCAAAACCA 540
QY 548 CTCGAAGGGCAGCAGCAATTTCTGTGTAAGAACATCAAGCAGGAGCAATTTGAGCGCG 607
Db 541 CTCGAAGGGCAGCAGCAATTTCTGTGAAGAACATCAAGCAGGAGCAATTTGAGCGCG 600
QY 608 GGAGATTGAGATTCAGAGCAAGACATGTGCTGCTGATTCTCACTCAGGAAGCTGCTCA 667
Db 601 GGAGATTGAGATTCAGAGCAAGACATGTGCTGCTGATTCTCACTCAGGAAGCTGCTCA 660
QY 668 GGGGAGAGCCCTTGGCTGTGCTAAATATAGTGGCTGTACACACATCAGAGCCAGAG 727
Db 661 GGGGAGAGCCCTTGGCTGTGCTAAATATAGTGGCTGTACACACATCAGAGCCAGAG 720

QY 728 AGCGGTGTGATGAGACACTCTGTGCCCTGGGGGCTCAGTGCCGCTGCTCTGTAA 787
|||||
Db 721 AGCGGTGTGATGAGACACTCTGTGCCCTGGGGGCTCAGTGCCGCTGCTCTGTAA 780
QY 788 CATCTACTCACTCAGAAATGAGTAGCTGCAGCAGCTGGCTGAGGCTGAGTTGAGTTT 847
|||||
Db 781 CATCTACTCACTCAGAAATGAGTAGCTGCAGCAGCTGGCTGAGGCTGAGTTGAGTTT 840
QY 848 CGCTTGGAAGGCGAGTCAGAGAT 872
|||||
Db 841 CGCTTGGAAGGCGAGTCAGAAAT 865

RESULT 13
BM468564 967 bp mRNA linear EST 05-FEB-2002
LOCUS BM468564
DEFINITION AGENCOURT_6475681 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5578170
5', mRNA sequence.
ACCESSION BM468564
VERSION BM468564.1 GI:18517606
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 967).
NIH-MGC <http://mgc.ncl.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12332 row: m column: 19
High quality sequence stop: 685.
Location/Qualifiers
1. 967
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5578170"
/clone_1id="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."

BASE COUNT 214 a 253 c 328 g 172 t
ORIGIN

Query Match 33.0%; Score 847; DB 13; Length 967;
Best Local Similarity 94.4%; Pred. No. 2.6e-143;
Matches 900; Conservative 0; Mismatches 50; Indels 3; Gaps 2;

QY 1 GCGCGGCGAGTGGAGCTCGAGCTGCTGCTTCTGCTCTGTGCGCACCGTGGCT 60
|||||
Db 3 GCGCGGCGAGTGGAGCTCGAGCTGCTGCTTCTGCTCTGTGCGCACCGTGGCT 62
QY 61 GTCGGGCTGCTTGGGCTGGCGAAGACAGACAGAGCGTGGGCGACAGCACTCAGAGCGA 120
|||||
Db 63 GTCGGGCTGCTTGGGCTGGCGAAGACAGACAGAGCGTGGGCGACAGCACTCAGAGCGA 122
QY 121 CGCAGCTCGACGACAGGCGCGACAGAGGTTGGGCGATCGCTGTGCGAGGCGCGCGC 180
|||||
Db 123 CGCAGCTCGACGACAGGCGCGCGACAGAGGTTGGGCGATCGCTGTGCGAGGCGCGCGC 182
QY 181 GGGCAGGCGGCG 240
|||||

Db 183 GGGCAGGCGGCGGCG 242
QY 241 GGGCGGCGGCGGGAATGTCATGCTGACGCGATGCCGCTGCCCGGGGTCGGGAGAGC 300
|||||
Db 243 GGGCGGCGGCGGGAATGTCATGCTGACGCGATGCCGCTGCCCGGGGTCGGGAGAGC 302
QY 301 TGAAGCAGCGCAGAGAGATCGAGAGCGCGGAGAGTACTCTTCATGGCCACCGTCACCA 360
|||||
Db 303 TGAAGCAGCGCAGAGAGATCGAGAGCGCGGAGAGTACTCTTCATGGCCACCGTCACCA 362
QY 361 AGCGCGCCAGAGCAATCCAGTTTGTGATGACATGACAGAGTTACCAATTCGCCA 420
|||||
Db 363 AGCGCGCCAGAGCAATCCAGTTTGTGATGACATGACAGAGTTACCAATTCGCCA 422
QY 421 CCAAACTGCGCGCAAGATCTTGTCTCGCTGATCTCAGAGTCCCTCCACTGACAGCTACA 480
|||||
Db 423 CCAAACTGCGCGCAAGATCTTGTCTCGCTGATCTCAGAGTCCCTCCACTGACAGCTACA 482
QY 481 GTTCAGCTGCATCTACACAGATAGCTGTGATGAGGTTTCTCCCGAGAGAGCAGC 540
|||||
Db 483 GTTCAGCTGCATCTACACAGATAGCTGTGATGAGGTTTCTCCCGAGAGAGCAGC 542
QY 541 AAACCACTCCAGAGCGCAGACGAATTTCTGTGAGAGCAATCAAGCAGGCGAGAATTG 600
|||||
Db 543 AAACCACTCCAGAGCGCAGACGAATTTCTGTGAGAGCAATCAAGCAGGCGAGAATTG 602
QY 601 GACGCCGGAGATGAGATTGACAGAGCAAGACATGCTGCTGTGATTTCACATCAGAAAC 660
|||||
Db 603 GACGCCGGAGATGAGATTGACAGAGCAAGACATGCTGCTGTGATTTCACATCAGAAAC 662
QY 661 GTGCTCAGGGGAGAGAGCCCTTGCTGTGCTAAATAGTGGGCTGTACACATCAGAG 720
|||||
Db 663 GTGCTCAGGGGAGAGAGCCCTTGCTGTGCTAAATAGTGGGCTGTACACATCAGAG 722
QY 721 CCAGACAGCGGCTGTGATTGAGACACTCTGCGCCCTGGGGCTCAGTGCCGCTGCTG 780
|||||
Db 723 CCAGACAGCGGCTGTGATTGAGACACTCTGCGCCCTGGGGCTCAGTGCCGCTGCTG 782
QY 781 CTGTACATCTACTCACTCAGATGAGTAGTGCAGCAGCTGGCTGAGGCTGAGTTG 840
|||||
Db 783 CTGT-AACATCTACTCACTCAGATGAGTAGTGCAGCAGCTGGCTGAGGCTGAGTTG 841
QY 841 CAGTGTGCTTGAAGGCGGAGTCAGAGATGAC--TTCTGTGTGTATTGACCGCTG 898
|||||
Db 842 CAGTGTGCTTGAAGGCGGAGTCAGAGATGAC--TTCTGTGTGTATTGACCGCTG 901
QY 899 TGTGACATGATGCGCTGGCGAGCGCAACATGATCTGTGATGATGGGGAGACT 951
|||||
Db 902 TGGGTGGAATGGAATGGGGGGCGCAAGCGCAACTGTATCCCTGGAATGAAT 954

RESULT 14
AL554850 858 bp mRNA linear EST 16-FEB-2001
LOCUS AL554850
DEFINITION AL554850 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI087YL18 5
prime, mRNA sequence.
ACCESSION AL554850
VERSION AL554850.1 GI:12896025
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 858)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. 858
source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1087YL18"
/clone_1lb="LTI_NFL006_PL2"
/tissue_type="Placenta"
/note="Vector: PCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the PCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 212 a 206 c 255 g 182 t 3 others
ORIGIN

Query Match 33.0%; Score 845.8; DB 9; Length 858;
Best Local Similarity 99.5%; Pred. No. 4.5e-143;
Matches 855; Conservative 3; Mismatches 0; Indels 1; Gaps 1;

QY 233 CGCTGCGCGCGCGCGCGCGGGAATGCGATGCGCGATGCGCGCGCGCGCGCGG 292
DB 1 CGCTGCGCGCGCGCGCGCGGGAATGCGATGCGCGATGCGCGCGCGCGCGCGG 59
QY 293 GGAGAGCTGAAGCAGGCGCAAGAGATGAGAGCGCGCGAGAGTACTCTTCATGGCCAC 352
DB 60 GGAGGCGCTGAAGCAGGCGCAAGAGATGAGAGCGCGCGAGAGTACTCTTCATGGCCAC 119
QY 353 CGTCAACCAAGCGCGCGCAAGAGAAATCCAGTTGCTGATGACATGAGAGTTTACCAA 412
DB 120 CGTCAACCAAGCGCGCGCAAGAGAAATCCAGTTGCTGATGACATGAGAGTTTACCAA 179
QY 413 ATTCGCCACCAAAACTGGCGGAAGATCTTGTCTGCTGCTGATCTCAGAGTCTCCACTGA 472
DB 180 ATTCGCCACCAAAACTGGCGGAAGATCTTGTCTGCTGCTGATCTCAGAGTCTCCACTGA 239
QY 473 CAGCTACAGTTCAGCTGCATCTCTACACAGATAGCTCTGATGATGAGTTTCTCCCGAGA 532
DB 240 CAGCTACAGTTCAGCTGCATCTCTACACAGATAGCTCTGATGATGAGTTTCTCCCGAGA 299
QY 533 GAAGCAGCAAAACCAACTCCCAAGGCGCAGCAATTTCTGTGAAGAACATCAAGCAGGC 592
DB 300 GAAGCAGCAAAACCAACTCCCAAGGCGCAGCAATTTCTGTGAAGAACATCAAGCAGGC 359
QY 593 AGAATTTGACGCGCGGAGATGAGTTGACAGCAAGACATGCTGCTGATTTCACT 652
DB 360 AGAATTTGACGCGCGGAGATGAGTTGACAGCAAGACATGCTGCTGATTTCACT 419
QY 653 CAGGAACGCTGCTCAGGGGAGAAAGCCCTGGCTGCTGCTAAATAGTGGGCTGTACACA 712
DB 420 CAGGAACGCTGCTCAGGGGAGAAAGCCCTGGCTGCTGCTAAATAGTGGGCTGTACACA 479
QY 713 CATCAACGCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGGGGCTCAGTGCCG 772
DB 480 CATCAACGCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGGGGCTCAGTGCCG 539
QY 773 CTGCTCTGCTGTAACATCTACTCAACTCAGAATGAAGTAGCTGCAGCACTGGCTGAGGC 832
DB 540 CTGCTCTGCTGTAACATCTACTCAACTCAGAATGAAGTAGCTGCAGCACTGGCTGAGGC 599
QY 833 TGGAGTTCAGTGTTCGCTTGAAGGCGAGTCAAGAGATGACTTCTGCTGCTGATTGA 892
DB 600 TGGAGTTCAGTGTTCGCTTGAAGGCGAGTCAAGAGATGACTTCTGCTGCTGATTGA 659
QY 893 CCGCTGCTGTAACATGATGGGTGCGAGCCCAACATGATCCTGATGATGGGAGACTT 952
DB 660 CCGCTGCTGTAACATGATGGGTGCGAGCCCAACATGATCCTGATGATGGGAGACTT 719
QY 953 AACCACTGGGTTTATAAGAGTATCAAAAGTGTTTAAGAGATCCGAGCATTTGTGA 1012
DB 720 AACCACTGGGTTTATAAGAGTATCAAAAGTGTTTAAGAGATCCGAGCATTTGTGA 779

QY 1013 AGAGAGCGTACTGCTGCTTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTCT 1072
DB 780 AGAGAGCGTACTGCTGCTTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTCT 839
QY 1073 TCCGGCCATGAAGCTCAAT 1091
DB 840 TCCGGCCATGAAGCTCAAT 858

RESULT 15
BQ059373
LOCUS

DEFINITION BQ059373 1021 bp mRNA linear EST 29-MAR-2002
AGENCOURT_7050082 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5815413
5', mRNA sequence.

ACCESSION BQ059373
VERSION BQ059373.1 GI:19818713
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1021)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Lou Staudt

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LNCM2070 row: J column: 22
High quality sequence stop: 688.

FEATURES
source 1.1021
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5815413"
/clone_1lb="NIH_MGC_99"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

BASE COUNT 284 a 245 c 201 g 291 t
ORIGIN

Query Match 32.9%; Score 844.2; DB 14; Length 1021;
Best Local Similarity 96.5%; Pred. No. 8.2e-143;
Matches 949; Conservative 0; Mismatches 23; Indels 11; Gaps 8;

QY 1260 ATTGTCTACATTACCGAATCGACCCCATCTGTGCTCTGACAGGCTGCATGATGGGTTTC 1319
DB 1 ATTGTCTACATTACCGAATCGACCCCATCTGTGCTCTGACAGGCTGCATGATGGGTTTC 60
QY 1320 AGGGTGTAAAGCTTAATGAAGTCATCCGGCAAGTCATGCTGTAATAACTTGCACAGGA 1379
DB 61 AGGGTGTAAAGCTTAATGAAGTCATCCGGCAAGTCATGCTGTAATAACTTGCACAGGA 120
QY 1380 AATAAGATGTAGTACACGGGAGCACTTGATGCGATGAATAACAGTTGTATGATGC 1439
DB 121 AATAAGATGTAGTACACGGGAGCACTTGATGCGATGAATAACAGTTGTATGATGC 180

QY 1440 AATATGGGCACTCCAAACACAGAAATCGATGTGACCAGCCTCCGCACTCCGAGCTGACG 1499
|||||
Db 181 AATATGGGCACTCCAAACACAGAAATCGATGTGACCAGCCTCCGCACTCCGAGCTGACG 240
QY 1500 TGGAGCGAGTAAGTCTCTCAGGTGAGCACCATGTCATCTGGCCAGATGGCAACGAGTTGTC 1559
|||||
Db 241 TGGAGCGAGTAAGTCTCTCAGGTGAGCACCATGTCATCTGGCCAGATGGCAACGAGTTGTC 300
QY 1560 CTCCCTGGCAGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTCCCACTTTGTCTG 1619
|||||
Db 301 CTCCCTGGCAGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTCCCACTTTGTCTG 360
QY 1620 TCCATCACAGCCACACACAGGCTTTGGCAGTGTAGAACTCTATAATGCACCCGAGGGG 1679
|||||
Db 361 TCCATCACAGCCACACACAGGCTTTGGCAGTGTAGAACTCTATAATGCACCCGAGGGG 420
QY 1680 CGATACAGCAGAGTGTACTGCTTCCCTAAAGAAATGATGATACGTTCGACGCTTG 1739
|||||
Db 421 CGATACAGCAGAGTGTACTGCTTCCCTAAAGAAATGATGATACGTTCGACGCTTG 480
QY 1740 CATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTG 1799
|||||
Db 481 CATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTG 540
QY 1800 GGACTCAACAAAAATGGGCCATTCAAACCTAATTAATTAACAGATTAATGGACCACTACTA 1859
|||||
Db 541 GGACTCAACAAAAATGGGCCATTCAAACCTAATTAATTAACAGATTAATGGACCACTACTA 600
QY 1860 CCAAGGACCAGTCCACTGAACACACACACTCTAAAGAAATTTTAAAGATAACTTTTA 1919
|||||
Db 601 CCAAGGACCAGTCCACTGAACACACACACTCTAAAGAAATTTTAAAGATAACTTTTA 660
QY 1920 TTTTCTCTTACTCTCTCTCTCTCTGATTTTCTCTATAATTTCAATCTGTTTTCAT 1979
|||||
Db 661 TTTTCTCTTACTCTCTCTCTCTCTGATTTTCTCTATAATTTCAATCTGTTTTCAT 720
QY 1980 CTCATATCCAAG-TTCTGCAGACACACAGAACTTCTCATGCTC-TTAAAGTAA 2037
|||||
Db 721 CTCATATCCAAGTCTCTGCAGACACACAGAACTTCTCTGCTCTTTAAATGAA 780
QY 2038 ATAGAAGTCA-GGGTCCCTCACTCTAGTCACTAAAGAAAGATTAC-TCCCCAGCCC 2095
|||||
Db 781 ATAGAAGTCAAGGGTCTCTCACTCTAGTCACTAAAGAAAGATTACTTCTCCAGCCC 840
QY 2096 AGAAGGTGATCTCTCTTTTACCATTCTGGGG--ACTTAGTCTTAA-TTAGGTACCT 2152
|||||
Db 841 GAAAAGGGGATCTTCTCTTACCATTTGGGGGAACTTAACTTAAGTTAGGTACCT 900
QY 2153 TATTAACAGGAAT-GCTAAGTACCTCTCTGTGAA--CAATCTGAATGTCTAAT 2208
|||||
Db 901 TATTAACAGGAATGTCTTAAGTACCTTCTCTGTGAAACAATCTGCATGGTCTAAT 960
QY 2209 CGCCTTAAAGAGCCCATTTCTT 2231
|||||
Db 961 CGCCTTAAAGAGCCCATTTCTT 983

Search completed: April 20, 2003, 17:23:39
Job time : 3309 secs

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2003, 21:15:39 ; Search time 88.6286 Seconds
(without alignments)
10595.712 Million cell updates/sec

Title: US-09-782-051-1_COPY_529_945

Perfect score: 417
Sequence: 1 gagagaagcagcaaccaac.....catgacctggtgatgaggg 417

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*

1:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	417	100.0	2563	19	AAV28617
2	417	100.0	3616	20	AAV73924
3	417	100.0	3634	23	ABV23195
4	417	100.0	3634	23	ABV29032
5	274.6	65.9	5030	23	AA576216
6	240	57.6	504	23	ABV44141
7	226.4	54.3	636	21	AAZ80766
8	225.8	54.1	312	23	ABV35309
9	225.8	54.1	435	23	ABV14215

10	208.8	50.1	1566	23	ABL03693	Drosophila melanog
11	191	45.8	1776	23	ABL29787	Drosophila melanog
12	191	45.8	5537	23	ABL29786	Drosophila melanog
13	178	42.7	721	21	AAF16004	Human prostate can
14	163.4	39.2	349	21	AAA44806	Human secreted exp
15	156.8	37.6	328	23	ABV05046	Human prostate exp
16	138.8	33.3	2211	22	AAC91228	Human S-adenosylho
17	137.2	32.9	2200	21	AAC77823	Human cancer assoc
18	137.2	32.9	2658	23	ABV23126	Human prostate exp
19	137.2	32.9	2658	23	ABV28967	Human prostate exp
20	130.2	31.2	656	21	AAF14585	Aspergillus oryzae
21	129.2	31.0	4406	23	ABL03692	Drosophila melanog
22	128.2	30.7	1000	22	AAF91420	Moraxella catarrha
23	128.2	30.7	1000	24	ABK37800	DNA sequence upstr
24	128.2	30.7	58909	22	AAF28543	Genomic fragment #
25	127.4	30.6	1592	23	ABL28157	Drosophila melanog
26	125.2	30.0	1865	21	AAC44037	Zea mays DNA fragm
27	118.2	28.3	1812	13	AAQ29419	Nicotiana tabacum
28	118.2	28.3	1812	17	AAQ29194	S-adenosylhomocyst
29	114.4	27.4	793	19	AAV28651	Ripening banana pu
30	114.4	27.4	1949	21	AAF07609	Fusarium venenatum
31	114.4	27.4	513445	22	AAI61373	Soybean 318013 reg
32	113	27.1	1396	22	AAF71864	Corynebacterium gl
33	113	27.1	1396	23	AAS96123	C. glutamicum gene
34	113	27.1	1422	22	AAH65801	C. glutamicum codin
35	113	27.1	1557	22	AAF71862	Corynebacterium gl
36	113	27.1	1557	23	AAS96121	C. glutamicum gene
37	113	27.1	1939	24	ABK52394	Sahh gene encoding
38	113	27.1	349980	22	AAH68526	C. glutamicum codin
39	106.8	25.6	708	22	AAH65802	C. glutamicum codin
40	102	24.5	1494	21	AAC46817	Arabidopsis thalia
41	101.8	24.4	1767	17	AAQ44513	Asparagus S-adenos
42	98.8	23.7	1728	21	AAC39836	Arabidopsis thalia
43	97.2	23.3	728	24	ABQ65540	Arabidopsis thalia
44	97.2	23.3	1727	21	AAC46315	Arabidopsis thalia
45	94	22.5	1941	17	AAQ44515	Arabidopsis S-aden

ALIGNMENTS

RESULT 1	AAV28617	standard; cDNA; 2563 BP.
ID	AAV28617;	
AC	AAV28617;	
XX		
DT	21-AUG-1998	(first entry)
DE	Nucleotide sequence of the ODD4b5.3 enzyme.	
XX		
KW	DD4b9.3 enzyme; S-adenosyl-L-homocysteine hydrolase-type activity;	
KW	AHCY-type activity; stimulation; inhibition; dendritic cell; cancer;	
KW	autoimmune disease; transplantation; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	3..1847
FT		/*tag= a
FT		/product= "DD4b5.3 enzyme"
XX		
PN	W09814562-A1.	
XX		
PD	09-APR-1998.	
XX		
PF	06-OCT-1997;	97WO-NZ00133.
XX		
PR	04-OCT-1996;	96NZ-0299507.
XX		
PA	(HART/) HART D N J.	
XX		
PI	Hart DNU;	

XX WPI: 1998-240074/21.
DR P-PSDB; AAW56097.
XX
PT New isolated S-adenosyl-L-homocysteine hydrolase enzyme - is used to
PT develop products which can be used in the treatment of e.g.
PT auto-immune disease, transplantations or cancers
XX
PS Claim 7; Fig 1; 33pp; English.
XX
CC This is the nucleotide sequence of the DD4b9.3 enzyme which has
CC S-adenosyl-L-homocysteine hydrolase (AHcy)-type activity. Its
CC products can be used to identify substances which have a stimulatory
CC or inhibitory effect on the enzyme activity. Such substances can be
CC used to modulate dendritic cell (DC) function and for immunomodulation.
CC They can be used in the treatment of e.g. autoimmune diseases,
CC transplantations or cancers. The products can also be used for
CC detection and diagnosis.
XX
SQ Sequence 2563 BP; 646 A; 604 C; 677 G; 636 T; 0 other;

Query Match 100.0%; Score 417; DB 19; Length 2563;
Best Local Similarity 100.0%; Pred. No. 2.2e-121;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAGAGAAGCAGCAACCACTCCAGGAGGAGAGCAATTTCTGTGAGAACAATCAAGC 60
DB 529 GAGAGAAGCAGCAACCACTCCAGGAGGAGAGCAATTTCTGTGAGAACAATCAAGC 588
OY 61 AGGAGAAATTTGACGCCGGGAGATGAGATTGACAGCAAGACATGTCCTCTGATT 120
DB 589 AGGAGAAATTTGACGCCGGGAGATGAGATTGACAGCAAGACATGTCCTCTGATT 648
OY 121 CACTCAGGAACGTCCTCAGGGGAGAGAGCCCTGCTGCTAAATAGTGGCTGTA 180
DB 649 CACTCAGGAACGTCCTCAGGGGAGAGAGCCCTGCTGCTAAATAGTGGCTGTA 708
OY 181 CACACATCACAGCCAGACAGCGGTGTTGATTGAGACACTCTGTGCCCTGGGGCTCAGT 240
DB 709 CACACATCACAGCCAGACAGCGGTGTTGATTGAGACACTCTGTGCCCTGGGGCTCAGT 768
OY 241 GCCGCTGCTGCTGTAACATCTACTCACTCAGATGAAGTAGCTGCAGCACTGGCTG 300
DB 769 GCCGCTGCTGCTGTAACATCTACTCACTCAGATGAAGTAGCTGCAGCACTGGCTG 828
OY 301 AGGCTGAGTTGAGTGTCTGCTTGAAGGGGAGTCAGAGATGACTTCTGTGTGTA 360
DB 829 AGGCTGAGTTGAGTGTCTGCTTGAAGGGGAGTCAGAGATGACTTCTGTGTGTA 888
OY 361 TTGACCGCTGTGTAACATGATGGTGGCAGGCCAATGATCTCTGATGGG 417
DB 889 TTGACCGCTGTGTAACATGATGGTGGCAGGCCAATGATCTCTGATGGG 945

RESULT 2
AAV73924
ID AAV73924 standard; DNA; 3616 BP.

XX AAV73924;

DT 04-MAR-1999 (first entry)

DE Human SAHH DNA #1.

KW S-adenosyl-5-homocysteine hydrolase; SAHH; human; drug screening;
KW treatment; infection; cancer; autoimmune disease; detection; diagnosis;
KW gene mapping; antisense; therapy; antagonist; immunoassay; ss.

OS Homo sapiens.

XX Key Location/Qualifiers
FH 56..1558
FT CDS /*tag= a

FT /product= "SAHH"
XX US5854023-A.
XX 29-DEC-1998.
XX 17-JUL-1997; 97US-0896005.
XX 17-JUL-1997; 97US-0896005.
XX 17-JUL-1997; 97US-0896005.
XX (INCYTE) INCYTE PHARM INC.
XX Corley NC, Hillman JL, Lai P, Shah P;
XX WPI: 1999-094906/08.
XX P-PSDB; AAW90061.
XX
PT Nucleic acid encoding human S-adenosyl-5-homocysteine hydrolase
PT for production of recombinant enzyme, useful for diagnosis,
PT treatment and prevention of cancers, infections and autoimmune
PT diseases
XX
XX Disclosure; Fig 1A-I; 40pp; English.

CC This sequence encodes a human S-adenosyl-5-homocysteine hydrolase (SAHH).
CC The SAHH protein can be used to generate specific antibodies and in drug
CC screening to identify specific binding agents. Antagonists of the
CC protein are used to treat or prevent a wide range of viral, bacterial,
CC fungal, parasitic, protozoal or helminthic infections, many cancers
CC (leukaemia, lymphoma or solid tumours), and many autoimmune diseases
CC (e.g. acquired immune deficiency syndrome, allergy, asthma, diabetes
CC mellitus, multiple sclerosis etc). All these conditions may be treated by
CC expressing antisense sequences, triple-forming agents or ribozymes
CC directed against the nucleic acid. The nucleic acid and its fragments can
CC be used as probes or primers for detecting and quantifying gene
CC expression, for diagnosis or monitoring of disease, to identify genetic
CC variations, mutations or polymorphisms, in gene mapping and as antisense
CC therapeutics. Antibodies are used directly as antagonists, indirectly to
CC deliver active agents to SAHH-expressing cells, to diagnose and monitor
CC diseases in standard immunoassays, in competitive drug screens and to
CC isolate the protein from natural sources.
XX

SQ Sequence 3616 BP; 1017 A; 782 C; 826 G; 991 T; 0 other;

Query Match 100.0%; Score 417; DB 20; Length 3616;
Best Local Similarity 100.0%; Pred. No. 2.6e-121;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAGAGAAGCAGCAACCACTCCAGGAGGAGAGCAATTTCTGTGAGAACAATCAAGC 60
DB 240 GAGAGAAGCAGCAACCACTCCAGGAGGAGAGCAATTTCTGTGAGAACAATCAAGC 299
OY 61 AGGAGAAATTTGACGCCGGGAGATGAGATTGACAGCAAGACATGTCCTGATT 120
DB 300 AGGAGAAATTTGACGCCGGGAGATGAGATTGACAGCAAGACATGTCCTGATT 359
OY 121 CACTCAGGAACGTCCTCAGGGGAGAGAGCCCTTGCTGCTAAATAGTGGCTGTA 180
DB 360 CACTCAGGAACGTCCTCAGGGGAGAGAGCCCTTGCTGCTAAATAGTGGCTGTA 419
OY 181 CACACATCACAGCCAGACAGCGGTGTTGATTGAGACACTCTGTGCCCTGGGGCTCAGT 240
DB 420 CACACATCACAGCCAGACAGCGGTGTTGATTGAGACACTCTGTGCCCTGGGGCTCAGT 479
OY 241 GCCGCTGCTGCTGTAACATCTACTCACTCAGATGAAGTAGCTGCAGCACTGGCTG 300
DB 480 GCCGCTGCTGCTGTAACATCTACTCACTCAGATGAAGTAGCTGCAGCACTGGCTG 539
OY 301 AGGCTGAGTTGAGTGTCTGCTTGAAGGGGAGTCAGAGATGACTTCTGTGTGTA 360
DB 540 AGGCTGAGTTGAGTGTCTGCTTGAAGGGGAGTCAGAGATGACTTCTGTGTGTA 599
OY 361 TTGACCGCTGTGTAACATGATGGTGGCAGGCCAATGATCTCTGATGGG 417

|||||
Db 600 TTGACCGCTGTGTGACATGATGGGTGGCAGGCCAACATGATCCTGGATGATGGG 656

RESULT 3.

ABV23195
ID ABV23195 standard; cDNA; 3634 BP.

XX AC ABV23195;

XX DT 16-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 23186.

XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.

OS Homo sapiens.

XX PN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US05171.

XX PR 17-FEB-2000; 2000US-183319P.

XX PR 16-MAR-2000; 2000US-189862P.

XX PR 25-MAY-2000; 2000US-207454P.

XX PR 09-JUN-2000; 2000US-211314P.

XX PR 18-JUL-2000; 2000US-219007P.

XX PR 13-DEC-2000; 2000US-255281P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE;

XX DR WPI; 2001-662795/76.

XX PS Claim 1; Page 4168-4169; 11750pp; English.

XX CC The invention relates to an isolated nucleic acid molecule (I) comprising

CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

CC specification or its complement. (I) is useful for:

CC (a) assessing whether a patient is afflicted with prostate cancer;

CC (b) monitoring the progression of prostate cancer in a patient;

CC (c) assessing the efficacy of a test compound to inhibit prostate

CC cancer in a patient;

CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer

CC in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;

CC (f) assessing the prostate cell carcinogenic potential of a compound;

CC (g) determining whether prostate cancer has metastasized in a patient;

CC (h) assessing the aggressiveness or indolence of prostate cancer in a

CC patient;

CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX SQ Sequence 3634 BP; 876 A; 966 C; 869 G; 918 T; 5 other;

Query Match 100.0%; Score 417; DB 23; Length 3634;

Best Local Similarity 100.0%; Pred. No. 2.6e-121;

Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAGAGCAGCAAACTCCAAAGGAGCAGCAATTTCTGTGAGAGACATCAAGC 60

Db 599 GAGAGAGCAGCAAACTCCAAAGGAGCAGCAATTTCTGTGAGAGACATCAAGC 658

QY 61 AGGAGAGATTTGAGCGCGGAGATGAGATTGACAGCAAGACATGTCGCTGATTT 120

Db 659 AGGAGAGATTTGAGCGCGGAGATGAGATTGACAGCAAGACATGTCGCTGATTT 718

QY 121 CACTCAGAAACGTGCTCAGGGGGAGAAAGCCCTTGCTGCTAAATAGTGGCTGTA 180

Db 719 CACTCAGAAACGTGCTCAGGGGGAGAAAGCCCTTGCTGCTAAATAGTGGCTGTA 778

QY 181 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTGTGCCCTGGGGCTCAGT 240

Db 779 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTGTGCCCTGGGGCTCAGT 838

QY 241 GCCGCTGCTGCTGTGTAACATCTCACTCAGAAATGAGTAGCTGCAGACTGGCTG 300

Db 839 GCCGCTGCTGCTGTGTAACATCTCACTCAGAAATGAGTAGCTGCAGACTGGCTG 898

QY 301 AGGCTGAGTGCAGTGTCCCTTGAAGGGCGAGTCAGAGATGACTTCTGTGCTGA 360

Db 899 AGGCTGAGTGCAGTGTCCCTTGAAGGGCGAGTCAGAGATGACTTCTGTGCTGA 958

QY 361 TTGACCGCTGTGTGACATGATGGGTGGCAGGCCAACATGATCCTGATGAGGG 417

Db 959 TTGACCGCTGTGTGACATGATGGGTGGCAGGCCAACATGATCCTGATGAGGG 1015

RESULT 4

ABV29032
ID ABV29032 standard; cDNA; 3634 BP.

XX AC ABV29032;

XX DT 16-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 29023.

XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.

OS Homo sapiens.

XX PN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US05171.

XX PR 17-FEB-2000; 2000US-183319P.

XX PR 16-MAR-2000; 2000US-189862P.

XX PR 25-MAY-2000; 2000US-207454P.

XX PR 09-JUN-2000; 2000US-211314P.

XX PR 18-JUL-2000; 2000US-219007P.

XX PR 13-DEC-2000; 2000US-255281P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE;

XX DR WPI; 2001-662795/76.

XX PS Claim 1; Page 6147-6148; 11750pp; English.

XX CC The invention relates to an isolated nucleic acid molecule (I) comprising

CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

CC specification or its complement. (I) is useful for:

CC (a) assessing whether a patient is afflicted with prostate cancer;

CC (b) monitoring the progression of prostate cancer in a patient;

CC (c) assessing the efficacy of a test compound to inhibit prostate

CC cancer in a patient;

CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer

CC in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;

CC (f) assessing the prostate cell carcinogenic potential of a compound;

CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX

Sequence 3634 BP; 876 A; 966 C; 869 G; 918 T; 5 other;

Query Match 100.0%; Score 417; DB 23; Length 3634;
Best Local Similarity 100.0%; Pred. No. 2.6e-121;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAGAGCAGCAACCACTCCAGGCGAGCAGCAATTTCTGTGAGAGACATCAAGC 60
DB 599 GAGAGAGCAGCAACCACTCCAGGCGAGCAGCAATTTCTGTGAGAGACATCAAGC 658
QY 61 AGGCAAGATTGGACGCCGGGAGATGAGATTGCAGAGCAAGACATGCTCTGATTT 120
DB 659 AGGCAAGATTGGACGCCGGGAGATGAGATTGCAGAGCAAGACATGCTCTGATTT 718
QY 121 CACTCAGGAACGTGCTCAGGGGAGAGCCCTGGCTGTGCTAAATAGTGGGCTGTA 180
DB 719 CACTCAGGAACGTGCTCAGGGGAGAGCCCTGGCTGTGCTAAATAGTGGGCTGTA 778
QY 181 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTGTGCTGGGGGCTCAGT 240
DB 779 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTGTGCTGGGGGCTCAGT 838
QY 241 GCCGCTGCTCTGCTTGTACATCTACTCACTCAGATGAGTAGCTGCAGCACTGGCTG 300
DB 839 GCCGCTGCTCTGCTTGTACATCTACTCACTCAGATGAGTAGCTGCAGCACTGGCTG 898
QY 301 AGGCTGAGTTCAGAGTGTTCGCTTGAAGCGCGAGTCAAGAGATGACTTCTGTGTGTA 360
DB 899 AGGCTGAGTTCAGAGTGTTCGCTTGAAGCGCGAGTCAAGAGATGACTTCTGTGTGTA 958
QY 361 TTGACCGCTGTGTGAACATGAGTGGGTGGCAGGCCAATGATCCTGATGATGGG 417
DB 959 TTGACCGCTGTGTGAACATGAGTGGGTGGCAGGCCAATGATCCTGATGATGGG 1015

RESULT 5

AAS76216/C
ID AAS76216 standard; cDNA; 5030 BP.

AC AAS76216;
DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #12020.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR P-PSDB; ABG12029.

PT New isolated polynucleotide and encoded polypeptides, useful in
diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess
PT biodiversity

PS Claim 1; SEQ ID NO 12020; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations,
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 5030 BP; 1432 A; 1137 C; 1178 G; 1283 T; 0 other;

Query Match 65.9%; Score 274.6; DB 23; Length 5030;
Best Local Similarity 78.7%; Pred. No. 3.3e-76;
Matches 328; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 1 GAGAGAGCAGCAACCACTCCAGGCGAGCAGCAATTTCTGTGAGAGACATCAAGC 60
DB 4489 GGGACAGCAGCAAGCAAGCACTTAAGGAGAGCAGTCTGTGTTAAGACATCAAC 4430
QY 61 AGGCAAGATTGGACGCCGGGAGATGAGATTGCAGACCAAGACATGCTGCTGATTT 120
DB 4429 AGGCAAGATTGGACAGAGAAATTGAATTCCTGAACAAGAAATGCCCTGATGATGG 4370
QY 121 CACTCAGGAACGTGCTCAGGGGAGAGCCCTGGCTGTGCTAAATAGTGGGCTGTA 180
DB 4369 CTTTGAGGAAGAGAGCTCAAGGAGAAAGCCCTTGGCTGAGGCCAAATCGTGGCTGCA 4310
QY 181 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTGTGCCCCGCGGCTCAGT 240
DB 4309 CACACATCACAGCTCAGACTGCTGTGCTTATGAACTCTGGGTGCGGCGCCAGT 4250
QY 241 GCCGCTGCTCTGCTTGAACATCTACTCACTCAGATGAAGTAGCTGCAGCACTGGCTG 300
DB 4249 GCCGATGGCTGCCCAACATCTATTCCACTCTCAATGAAGTGGCTGCTGCTCAGCAG 4190
QY 301 AGGCTGAGTTCAGAGTGTTCGCTTGAAGGCGAGTCAAGAGATGACTTCTGTGTGTA 360
DB 4189 AAGTGGATTTCCTGTTTTCCTGGAAGGAGAGTCAAGATGACTTTGGTGTGTA 4130
QY 361 TTGACCGCTGTGTGAACATGAGTGGGTGGCAGGCCAATGATCCTGATGATGGG 417
DB 4129 TCGATAGATGTGTGAATGTGAGGGGCTGGCAGCCCAACATGATCTTGATGATGAG 4073

RESULT 6

ABV44141/C
ID ABV44141 standard; cDNA; 504 BP.

AC ABV44141;

DT 16-SEP-2002 (first entry)

DE Human prostate expression marker cDNA 44132.

KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

```

KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX
PS Claim 1; Page 8770; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 504 BP; 117 A; 150 C; 116 G; 121 T; 0 other;

Query Match          57.6%; Score 240; DB 23; Length 504;
Best Local Similarity 100.0%; Pred. No. 1e-65;
Matches 240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 178 GTACACACATCACAGCCCAAGACGCGGTGTGATGTAGACACATCTGCCCCGCGGCTC 237
   |||||||
DB 502 GTACACACATCACAGCCCAAGACGCGGTGTGATGTAGACACATCTGCCCCGCGGCTC 443

OY 238 AGTCCCGCTGCTGCTGTGTACATCTACTCAACTCAGAAATGAGTAGCTGCAGCACTGG 297
   |||||||
DB 442 AGTCCCGCTGCTGCTGTGTACATCTACTCAACTCAGAAATGAGTAGCTGCAGCACTGG 383

OY 298 CTGAGCGCTGAGTGTGACAGTGTGCGCTTGGAAGGGCGAGTCAGAAATGACTTCTGTGGT 357
   |||||||
DB 382 CTGAGCGCTGAGTGTGACAGTGTGCGCTTGGAAGGGCGAGTCAGAAATGACTTCTGTGGT 323

OY 358 GTATTGACCGCTGTGTGAACATGAGATGGGTGGCAGGCCAACATGATCCTGGATGATGGGG 417
   |||||||
DB 322 GTATTGACCGCTGTGTGAACATGAGATGGGTGGCAGGCCAACATGATCCTGGATGATGGGG 263

RESULT 7
AAZ80766/c
ID AAZ80766 standard; cDNA; 636 BP.
XX
AC AAZ80766;
```

```

XX
DT 07-APR-2000 (first entry)
XX
DE Human colon cancer cell line SW480 cDNA clone SEQ ID NO:850.
XX
KW Human; gene expression product; diagnosis; tumour; colon cancer;
KW colorectal adenocarcinoma; cell line SW480; cell proliferation;
KW cytostatic; sarcoma; breast cancer; neoplasia; dysplasia;
KW hyperplasia; ds.
XX
OS Homo sapiens.
XX
PN WO9964576-A2.
XX
PD 16-DEC-1999.
XX
PF 09-JUN-1999; 99WO-IB01062.
XX
PR 10-JUN-1998; 98US-0088801.
XX
PA (FARB ) BAYER CORP.
XX
PI Endege WO, Steilmann KE, Astle JH, Burgess CC, Bushnell SE;
PI Carroll E, Catlino TJ, Dertl A, Ford DM, Lewis ME, Monahan JE;
PI Schlegel R;
XX
DR WPI; 2000-087220/07.
XX
PT Novel nucleic acids, used to develop products for the diagnosis and
PT treatment of disorders involving unwanted cell proliferation,
PT particularly cancers, especially colon cancer -
XX
PS Claim 15; Page 469; 469pp; English.
XX
CC AAZ79917 to AAZ80766 represent double stranded cDNA clones isolated from
CC the human colorectal adenocarcinoma (colon cancer) cell line SW480. The
CC cDNA clones can be used to generate antisense oligonucleotides which
CC can be used for antisense therapy. Methods and products from the present
CC invention can be used for identifying and/or classifying cancerous cells
CC present in a human tumour, particularly in solid tumours, e.g.
CC carcinomas and sarcomas, e.g. breast or colon cancers. The cDNA clones
CC can be used for developing agents for the diagnosis and treatment of
CC disorders involving unwanted cell proliferation, such as neoplasia,
CC dysplasia or hyperplasia.
XX
SQ Sequence 636 BP; 156 A; 166 C; 137 G; 147 T; 30 other;

Query Match          54.3%; Score 226.4; DB 21; Length 636;
Best Local Similarity 99.2%; Pred. No. 2.3e-61;
Matches 238; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 179 TACACACATCACAGCCCAAGACAGC-GGTGTGATGTAGACACATCTGTGCCCTGCGGCTC 237
   |||||||
DB 420 TACACACATCACAGCCCAAGACAGCGGTGTGATGTAGACACATCTGTGCCCTGCGGCTC 361

OY 238 AGTCCCGCTGCTGCTGTGTACATCTACTCAACTCAGAAATGAGTAGCTGCAGCACTGG 297
   |||||||
DB 360 AGTCCCGCTGCTGCTGTGTACATCTACTCAACTCAGAAATGAGTAGCTGCAGCACTGG 301

OY 298 CTGAGCGCTGAGTGTGACAGTGTGCGCTTGGAAGGGCGAGTCAGAAATGACTTCTGTGGT 357
   |||||||
DB 300 CTGAGCGCTGAGTGTGACAGTGTGCGCTTGGAAGGGCGAGTCAGAAATGACTTCTGTGGT 241

OY 358 GTATTGACCGCTGTGTGAACATGAGATGGGTGGCAGGCCAACATGATCCTGGATGATGGGG 417
   |||||||
DB 240 GTATTGACCGCTGTGTGAACATGAGATGGGTGGCAGGCCAACATGATCCTGGATGATGGGG 181

RESULT 8
ABV35309
ID ABV35309 standard; cDNA; 312 BP.
XX
AC ABV35309;
```

XX	16-SEP-2002	(first entry)
XX	Human prostate expression marker	CDNA 35300.
DE	Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;	
XX	pharmacogenomic marker; gene; ss.	
KW	Homo sapiens.	
XX	WO200160860-A2.	
PN	23-AUG-2001.	
XX	20-FEB-2001; 2001WO-US05171.	
PF	17-FEB-2000; 2000US-183319P.	
XX	16-MAR-2000; 2000US-189862P.	
PR	25-MAY-2000; 2000US-207454P.	
PR	09-JUN-2000; 2000US-211314P.	
PR	18-JUL-2000; 2000US-219007P.	
PR	13-DEC-2000; 2000US-255281P.	
XX	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.	
PA	Schlegel R, Endege WO, Monahan JE;	
PI	WPI; 2001-662795/76.	
XX	Novel isolated nucleic acid molecule associated with cancerous state of	
PT	prostate cells and correlating with presence of prostate cancer, useful	
PT	for detecting presence of prostate cancer, stage of prostate cancer -	
XX	Claim 1; Page 7360-7361; 11750pp; English.	
XX	The invention relates to an isolated nucleic acid molecule (I) comprising	
CC	a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the	
CC	specification or its complement. (I) is useful for:	
CC	(a) assessing whether a patient is afflicted with prostate cancer;	
CC	(b) monitoring the progression of prostate cancer in a patient;	
CC	(c) assessing the efficacy of a test compound to inhibit prostate	
CC	cancer in a patient;	
CC	(d) assessing the efficacy of a therapy for inhibiting prostate cancer	
CC	in a patient;	
CC	(e) selecting a composition for inhibiting prostate cancer in a patient;	
CC	(f) assessing the prostate cell carcinogenic potential of a compound;	
CC	(g) determining whether prostate cancer has metastasized in a patient;	
CC	(h) assessing the aggressiveness or indolence of prostate cancer in a	
CC	patient;	
CC	(I) is also useful as a pharmacodynamic or pharmacogenomic marker.	
XX	Sequence 312 BP; 69 A; 71 C; 102 G; 70 T; 0 other;	
XX	Query Match 54.1%; Score 225.8; DB 23; Length 312;	
XX	Best Local Similarity 98.8%; Pred. No. 2.6e-61;	
XX	Matches 238; Conservative 0; Mismatches 2; Indels 1; Gaps 1;	
QY	178 GTACACACATCACAGCCAGACAGCGGTGTTGATTGAGACACTCTGTGCCCTGGGGCTC	237
DB		
DB	53 GTACACACATCACAGCCAGACAGCGGTGTTGATTGAGACACTCTGTGCCCTGGGGCTC	112
QY	238 AGTCCGCTGCTGCTTGTACATCTACTCACTCAGATGAGTAGCTGCAGCACTGG	297
DB		
DB	113 AGTGTGCTGCTGCTGTTGTAACATCTACTCACTCAGATGAGTAGCTGCAGCACTGG	172
QY	298 CTGAGGCTGAGTGCAGTGTTCGCTTGGAGGGCGAGTCAGAGATG-ACCTTCTGGTGG	356
DB		
DB	173 CTGAGGCTGAGTGCAGTGTTCGCTTGGAGGGCGAGTCAGAGATGAGTGAATCTTCTGGTGG	232
QY	357 TGTATTGACCGCTGTGTGAACATGATGGGTGGCAGGCCAACATGATCCTGGATGATGGG	416
DB		
DB	233 GGTATTGACCGCTGTGTGAACATGATGGGTGGCAGGCCAACATGATCCTGGATGATGGG	292

QY	417	G	417
Db	293	G	293

RESULT 9
ABV14215
ID ABV14215 standard; cDNA; 435 BP.

AC ABV14215;

DT 13-SEP-2002 (first entry)

DE Human prostate expression marker cDNA 14206.

KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

11-11-66

XX
XX
E000001600000-00

XX
2007-06-28

XX
XX
3
FEB
2001
000170
2206177

XX 17-FEB-2000 2000TC-163310T

PR 16-MAR-2000; 2000US-189862P.
35-MAR-2000; 2000US-207454D.

PR 09-JUN-2000; 2000US-211314P.
18-JUN-2000; 2000US-211314P.
18-JUN-2000; 2000US-211314P.

PR 13-DEC-2000; 2000US-255281P.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX

PI Schlegel R, Endege WO, Monahan JE;
 xx

DR WPT; 2001-662795/76.
XX

PT Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer [abstract]

PT for detecting presence of prostate cancer, stage of prostate cancer
XY

PS Claim 1; page 2369-2370; 11750pp; .English
XX

CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62313) of the

cc specification of its complement. (1) is useful for:
cc (a) assessing whether a patient is afflicted with prostate cancer;

(b) monitoring the progression of prostate cancer in a patient;

CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer

CC In a patient;
CC (e) selecting a common +11 on prostate cancer of a patient

(f) assessing the prostate cell carcinogenic potential of a compound;

CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient:

(1) is also useful as a pharmacodynamic or pharmacogenomic marker.

sequence 435 bp; 10/ A; 86 C; 133 G; 109 T; 0 other;

Query Match	54.18;	Score 225.8;	DB 23;	Length 435;
Best Local Similarity	98.88;	Pred NO	3a-61;	

Matches 238; conservative 0; mismatches 2; indels 1; gaps 1;

1/8 GTACACACATCACAGCCACGACAGCGGTGTGATTGAGACACATCTGTGCCCTGGGGCTC 23/

DB 3 GTACACACATCACAAGCCCAAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGGGGCTC 62

238 AGGCGGCGGCTGCTGCTGACATCTACTCAGATGAGTAGCTGCAGCAGCTGG 297

Db	63	AGTGTGCGCTGGTCTGCTTGTAACATCTACTCAACTCAGAAATGAAGTAGCTGCAGCACACTGG	122
OY	298	CTGAGGCGCTGGAGTTGCAGTGTTCGCTTGGAAAGGGCGAGTCAGAAGATG-ACTTCTGGTGG	356
Db	123	CTGAGGCGCTGGAGTTGCAGTGTTCGCTTGGAAAGGGCGAGTCAGAAGATGAACTTCTGCTGGTGG	182
OY	357	TGTATTGACCGCGCTGTGTGAACATGCGATGGGTGGCAGGCCAACATGATCCTGGATGATGGG	416
Db	183	GGTATTGACCGCGCTGTGTGAACATGCGATGGGTGGCAGGCCAACATGATCCTGGATGATGGG	242
OY	417	G 417	
Db	243	G 243	

RESULT 10
ABL03693
ID ABL03693 standard; cDNA; 1566 BP.

DT	26-MAR-2002 (first entry)
XX	
DE	Drosophila melanogaster expressed polynucleotide SEQ ID NO 5561.

Drosophila; developmental biology; cell signalling; insecticide;
pharmaceutical; gene; ss.

05 *Drosophila melanogaster*.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

23-MAR-2000; 2000US-191637P.

XX

PI Venter JC, Adams M, LI PWD, Myers EW;

DR WPI; 2001-656860/75.

DR P-PSDB; ABB59590.

PT New Isolated nucleic acid detection reagent for detecting 1000 or more

PT Interactions -

PS Claim 1; SEQ ID NO 5561; 21pp + Sequence Listing; English.

The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from *Drosophila*. The invention is useful in developmental biology and in elucidating cell signaling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB116176-AB130511), expressed DNA sequences (AB101840-AB116175) and the encoded proteins (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

Sequence 1566 BP; 360 A; 438 C; 481 G; 287 T; 0 other;

Query Match	50.1%;	Score 208.8;	DB 23;	Length 1566;
Best Local Similarity	70.28;	Pred. No. 1.3e-55;		
Matches 295; Conservative	0;	Mismatches 122;	Indels 3;	Gaps 1;

```

Qy      1 GAGGAAGCAGCAACCAACTCCAAAGGCGAGCAGCAATTTCTGTGTGAAGAACAT---CA 57
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    245 GGGAGAGGCTCCAAAAGAACTCCAAAGCGAGCTCCGATTTCGTGTGAAGGAACATAGCG 304

```

QY	58	AGCAGCGAGAATTTGGACGCCGCGAGATTGAGATTGCAGAGCAAGACATGTCTGCTTGA	117
Dd	305	CCACGATGCATTTGGGCGCAGGGAGATCGAGATTGCGGAGCAGGAGATGCCGGGCATCA	364
QY	118	TTTACTCAGGAAACGTGCTCAGGGGGGAGACCCTTGCGTGTCTAAAATAGTGGGCT	177
Dd	365	TAGCGCTGAGAAAGCGGGCGCGAGGACAACCTCTGAAGGATGCCAAGATCGTGGGTT	424
QY	178	GTCACACATCACAGCCCGACAGCAGCGGTGTGATTGAGACACTGTGCCCCGCGGGCTC	237
Dd	425	GCACACACATCAACGCCCGAGACTGCCGTGTCATCGAGACGCTGTGTAACCTGGCGCCA	484
QY	238	AGTGCCTGTGCTGCTTGTAACATCTACTCAACTCAGAATGAAGTAGCTGCACACTGG	297
Dd	485	GCGTGCCTGGGCGCCGCTGCACACATCTACTCCACGCAAGACGAAGTGGCTGCTTTGG	544
QY	298	CTGAGGCTGGAGTTGCAGTGTTCGCTTGGAAAGGCGAGTCAGAAGATGACTTCTGGTGGT	357
Dd	545	CCGAATCCGGAATCCCAGATCTTTGCTCTGGCGCGCGGAGACGGAAGAGACTTCTGGTGGT	604
QY	358	GTAATTGACCGCTGTGTGAACATGGATGGGTGGCAGGCCAACATGATCCTGGATGATGGCG	417
Dd	605	GCAATGATCGCTGCGTTAACGCCGAGAACTGGCAGCCCAAACATGATCCTGGACGATGGCG	664

RESULT 11	
ABL29787	
ID	ABL29787 standard; DNA; 1776 BP.

AC ABL29787;

DT- 26-MAR-2002 (first entry)

DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 40834..

KW Drosophila; developmental biology; cell signalling; insecticide;

XX

✕

PN WO200171042

PD 27-SEP-2001.

XX

XX

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP. NY.

PI Venter JC, Adams M, Ll PWD, Myers EW;

DR WPI; 2001-656860/75.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from *Drosophila* and for elucidating cell signaling and cell-cell
PT interactions -
XX

PS Claim 1; SEQ ID NO 40834; 21pp + Sequence Listing; English.

The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from *Drosophila*. The invention is useful in developmental biology and in elucidating cell signaling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins.

The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1776 BP; 437 A; 432 C; 479 G; 428 T; 0 other;
SQ
Query Match 45.8%; Score 191; DB 23; Length 1776;
Best Local Similarity 66.3%; Pred. No. 5.8e-50;
Matches 275; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
QY 3 GAGAGCAGCAACCACTCCAGAGGCGAGCAGCAATTTCTGTGTAAGACATCAAGCAG 62
DB 361 GACAACCATCAGCGCACTCCCGGAGGCGACTGCTTTGCTGAAGACATCTCGAAG 420
QY 63 GCAGATTGGAGCGCGGAGATTGAGATTGACAGACATGCTGCTGTGATTCA 122
DB 421 AGTCATTGGAGGCGGAGATGAGATGCGCGAGTCGAGATGCGGCATCATGACT 480
QY 123 CTCAGGAACGTGCTCAGGGGAGAAAGCCCTTGCTGCTAAATAGTGGCTGTACA 182
DB 481 CTGAGGAAAGAGCGAAGATGAGAAAGCCCTAAAGGTGCCAATATCGTCGATGCAAC 540
QY 183 CACATCACAGCCAGACAGCGGTGTGATTGAGACACTGCTGCCCTGGGGCTCAGTGC 242
DB 541 CACGTCAATGCTCAGTCGCGAGTGTGATCGAGACCCCTGCTCCAATGCGGCCACAGTT 600
QY 243 CGCTGCTGCTGTGTAACATCTACTCAACTCAGATGAGTACTGCTGCTGCTGAT 302
DB 601 CGCTGGGCTGCTGCAACATTTATTCACACAAAACGAGTTGCCCGCTCTGGCAGAG 660
QY 303 GCTGAGTGTGAGTGTGCTTGGAGGGCGAGTCAAGATGACTTCTGCTGTGATT 362
DB 661 GCGGGAATTCGATCTTCGCTGCGCGGAGAGACGGAGAGAGTCTGTGCTGTG 720
QY 363 GACCGCTGTGTAACATGATGGGTGCGAGGCCAATGATCTGTGATGAGG 417
DB 721 GACAGGGCCATCTACTCCGACGGCTGGCAGCCGAACTAATCTGACGAGCGG 775
RESULT 12
ABL29786 standard; DNA; 5537 BP.
ID ABL29786;
XX ABL29786;
AC
XX 26-MAR-2002 (first entry)
DT
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 40831.
DE
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 40831; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 5537 BP; 1632 A; 1054 C; 1158 G; 1693 T; 0 other;
Query Match 45.8%; Score 191; DB 23; Length 5537;
Best Local Similarity 66.3%; Pred. No. 9.6e-50;
Matches 275; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
QY 3 GAGAGCAGCAACCACTCCAGAGGCGAGCAATTTCTGTGTAAGACATCAAGCAG 62
DB 2919 GACAACCATCAGCGCACTCCCGGAGGCGACTGCTTTGCTGAAGACATCTCGAAG 2978
QY 63 GCAGATTGGAGCGCGGAGATTGAGATTGACAGCAAGACATGCTGCTGTGATTCA 122
DB 2979 AGTCATTGGAGGCGGAGATGAGATGCGCGAGTGGAGATGCGGGCATCATGACT 3038
QY 123 CTCAGGAACGTGCTCAGGGGAGAAAGCCCTTGCTGCTAAATAGTGGCTGTACA 182
DB 3039 CTGAGGAAAGAGCGAAGATGAGAAAGCCCTAAAGGTGCCAATATGCTCGATGCAAC 3098
QY 183 CACATCACAGCCAGACAGCGGTGTGATTGAGACACTGCTGCCCTGGGGCTCAGTGC 242
DB 3099 CACGTCAATGCTCAGTCGCGAGTGTGATCGAGACCCCTGCTCCAATGCGGCCACAGTT 3158
QY 243 CGCTGCTGCTGTGTAACATCTACTCAACTCAGATGAGTACTGCTGCTGATG 302
DB 3159 CGCTGGGCTGCTGCAACATTTATTCACACAAAACGAGTTGCCCGCTCTGGCAGAG 3218
QY 303 GCTGAGTGTGAGTGTGCTTGGAGGGCGAGTCAAGATGACTTCTGCTGTGATT 362
DB 3219 GCGGGAATTCGATCTTCGCTGCGCGGAGAGACGGAGAGAGTCTGTGCTGTG 3278
QY 363 GACCGCTGTGTAACATGATGGGTGCGAGGCCAATGATCTGTGATGAGG 417
DB 3279 GACAGGGCCATCTACTCCGACGGCTGGCAGCCGAACTAATCTGACGAGCGG 3333
RESULT 13
AAF16004 standard; cDNA; 721 BP.
ID AAF16004
XX AAF16004;
AC AAF16004;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:439.
XX
KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
KW vulnery; gastrointestinal; nephrotropic; antiinfective; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW wound; infectious disease; ss.
XX
XX Homo sapiens.
XX
PN WO200055174-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US05988.
XX
PR 12-MAR-1999; 99US-0124270.
XX

PA	(HUMA-)	HUMAN GENOME SCI INC.
PA	(ROSE/)	ROSEN C A.
XX	Rosen CA,	Ruben SM;
XX	PI	
XX	DR	WPI; 2000-587513/55.
XX	DR	P-PsDB; AAB56801.
PT	Prostate cancer associated gene sequences, referred to as prostate	
PT	cancer antigens, useful for treatment, prevention, and diagnosis of	
PT	disorders such as prostate cancer -	
PS	claim 1; page 917-918; 2338pp; English.	
XX		
CC	AAFI5566 to AAFI6505 encode the human prostate cancer associated	
CC	proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.	
CC	The prostate cancer antigens can have neuroprotective, cytostatic,	
CC	cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,	
CC	nephrotoxic, antinfecive, gynaecological and antibacterial activities,	
CC	and can be used in gene therapy. The prostate cancer antigen	
CC	polynucleotides may be used for detection of prostate cancer, chromosome	
CC	identification, as chromosome markers, and for numerous other diagnostic	
CC	or research purposes. The prostate cancer antigens may be used to treat	
CC	disorders such as neural, immune, muscular, reproductive,	
CC	gastrointestinal, pulmonary, cardiovascular, renal, and proliferative	
CC	disorders, wounds, and infectious diseases. AAFI6506 to AAFI6514 to	
CC	AAB57303 represent sequences used in the exemplification of the present	
CC	invention.	
XX		
SQ	Sequence 721 BP; 157 A; 195 C; 243 G; 119 T; 7 other;	
	Query Match	42.7%; Score 178; DB 21; Length 721;
	Best Local Similarity	94.2%; Pred. No. 5e-46;
	Matches 178; Conservative	5; Mismatches 6; Indels 0; Gaps 0;
OY	1	GAGAGAAGCAGCAACCACTCCAAAGGGGAGAGCAATTTCTGTGTGAAGACATCAAGC 60
DB	532	GAGAGAAGCAGCAACCACTCCAAAGGGGAGAGCAATTTCTGTGTGAAGACATCAAGC 591
OY	61	AGGCAGATTGTGAGCGCCGGAGATTGAGATTGAGAGCAAGACATGTCCTGTGATT 120
DB	592	AGGCAGATTGTGAGCGCCGGAGATTGAGATTGAGAGCAAGACATGTCCTGTGATT 651
OY	121	CACATCAGGAACGTGCTCAGGGGGAGAGACCCTTGCTGCTTAATAAGTGGGCTGA 180
DB	652	CACATCAGGAACGTGCTCAGGGGGAGARAAGCCCTTGCTGCTTAATAAKKGGGYTTGA 711
OY	181	CACACATCA.189
DB	712	CACATTTACA 720
RESULT 14		
ID	AAAA44806	standard; cDNA; 349 BP.
AC	AAAA44806;	
XX	21-AUG-2000	(first entry)
DE	Human secreted expressed sequence tag SEQ ID NO:1381.	
XX		
KM	Human; mouse; chicken; rat; secreted expressed sequence tag; SEST;	
KM	expressed sequence tag; EST; probe; chemotactic; proliferative;	
KM	immunomodulatory; haematopoietic; chemokineti; analgesic; haemostatic;	
KM	thrombolytic; antiinflammatory; cytosstatic; antibacterial; antifungal;	
KM	antiviral; antidiabetic; antiasthmatic; vulnerary; antiparkinsonian;	
KM	antiulcer; osteopathic; neuroprotective; nootropic; antipsoriatic;	
KM	cerebroprotective; anticonvulsant; antidepressant; gene therapy;	
KM	vaccine; autoimmune disorder; multiple sclerosis; allergic condition;	
KM	insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;	
KM	lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;	
KM	central nervous system disorder; Alzheimer's disease; stroke;	

```

KW Parkinson's disease; Huntington's disease; coagulation disorder;
KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease;
KW tumour; infection; depression; psoriasis; ss.
XX
OS Homo sapiens.
XX
PN WO200021991-A1.
XX
PD 20-APR-2000.
XX
PF 15-OCT-1999; 99WO-US24206.
XX
PR 15-OCT-1998; 98US-0104436.
XX
PA (GENY ) GENETICS INST INC.
XX
PI Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C,
PI Merberg D, Treacy M, Bowman MR;
XX
DR WPI; 2000-317938/27.
XX
PT Isolated polynucleotides, and encoded proteins, comprising secreted
PT expressed sequence tags (SESTs), useful for treating various disorders
PT such as autoimmune, infectious, and central nervous system disorders -
XX
PS Claim 1; Page 559; 803pp; English.
XX
CC AAA43426 to AAA45925 represent specifically claimed secreted expressed
CC sequence tags (SESTs), isolated from human, mouse, chicken and rat
CC tissue sources. The SESTs can have a range of activities depending on
CC the tissues they were isolated from. The activities include:
CC chemotactic; proliferative; immunomodulatory; haematopoietic;
CC chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory;
CC cytostatic; antibacterial; antifungal; antiviral; antidiabetic;
CC antiasthmatic; vulnery; antitumor; osteopathic; neuroprotective;
CC neurotropic; antiparkinsonian; antipsoriatic; cerebroprotective;
CC anticonvulsant; and antidepressant. The SESTs can be used for gene
CC therapy and in vaccines. The SESTs are useful as probes for the
CC identification and isolation of full-length cDNAs and genomic DNA
CC molecules which correspond to the SESTs. Proteins encoded by the SESTs
CC are useful in assays for determining biological activity and raising
CC antibodies. They may be useful for treatment of autoimmune disorders
CC (multiple sclerosis, insulin dependent diabetes), allergic conditions
CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,
CC osteoporosis, osteoarthritis, central nervous system disorders
CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation
CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's
CC disease), tumours, bacterial, fungal or viral infections, depression and
CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given
CC in the exemplification of the present invention.
XX
SQ Sequence 349 BP; 98 A; 91 C; 87 G; 73 T; 0 other;
XX
Query Match 39.2%; Score 163.4; DB 21; Length 349;
Best Local Similarity 99.4%; Pred. No. 1.5e-41;
Matches 164; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GAGAGAAGCAGCAAAACCACTCCAAAGGCGAGCAGCAATTTCTGTGTGAAGACATCAAGC 60
DB 178 GAGAGAAGCAGCAAAACCACTCCAAAGGCGAGCAGCAATTTCTGTGTGAAGACATCAAGC 237
QY 61 AGCAGAAATTTGAGCGCGGGAGATTGAGATTGCAGACCAAGACATGTCTGCTGATTT 120
DB 238 AGCAGAAATTTGAGCGCGGGAGATTGAGATTGCAGACCAAGACATGTCTGCTGATTT 297
QY 121 CACTCAGGAAACGTCGTCAGGGGGAGAGAGCCCTTGCGTGTGCTA 165
DB 298 CACTCAGGAAACGTCGTCAGGGGGAGAGAGCCCTTGCGTGTGCTCA 342

```

XX AC ABV05046;
XX 13-SEP-2002 (first entry)
DT XX
XX
DE Human prostate expression marker CDNA 5037.
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
OS Homo sapiens.
XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US05171.
XX
XX 17-FEB-2000; 2000US-183319P.
XX 16-MAR-2000; 2000US-189862P.
XX 25-MAY-2000; 2000US-207454P.
XX 09-JUN-2000; 2000US-211314P.
XX 18-JUL-2000; 2000US-219007P.
XX 13-DEC-2000; 2000US-255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX
XX Claim 1; Page 857; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX Sequence 328 BP; 77 A; 67 C; 102 G; 77 T; 5 other;
SQ
Query Match 37.6%; Score 156.8; DB 23; Length 328;
Best Local Similarity 89.6%; Pred. No. 1.8e-39;
Matches 225; Conservative 0; Mismatches 15; Indels 11; Gaps 5;
QY 178 GTACACATCATCAGCCAGACA-GCGGTGTTGATTGAGACACTCTGTGCCCTGGGGCT 236
DB 19 GTAGACATCATCAGCCAGACAGCGGTGTTGATTGAGACACTCTGTGCCCTGGGGCT 78
QY 237 CAGTGGCGTGTCTGCTGTGAACATCACTCACTCAGATGAAGTAGCTGCAGCACTG 296
DB 79 CAGTGTGCTGTCTGCTGTGAACATCACTCACTCAGATGAAGTAGCTGCAGCACTG 138
QY 297 GCTGAGGCTGAGTTCAGTGTGCT--GGAAGGGCGAGTCAGAGATGACTTC---- 350
DB 139 GCTGAGGCTGAGTTCAGTGTGCT--GGAAGGGCGAGTCAGAGATGACTTC---- 350
QY 351 TGGTGTGTTATGACCGCTGTG--AACATGATGGGTG--CAGGCCAATGATCTCT 406

DB 199 TGGTGTGTTATGACCGCTGTGTTGAGACATTGGATGGGTGGCAGGCCAATGATCTCT 258
QY 407 GGATGATGGGG 417
DB 259 GGATGATGGGG 269

Search completed: April 21, 2003, 00:05:26
Job time : 97.6286 secs

TITLE Nagasawa, H. and Little, J.B.
Increased ultraviolet sensitivity and chromosomal instability related to p53 function in the xeroderma pigmentosum variant
JOURNAL Cancer Res. 59 (5), 1102-1108 (1999)
MEDLINE 99168517
PUBMED 10070969
REFERENCE 2 (bases 1 to 2258)
AUTHORS Volpe, J.P.G., McDowell, M., Jostes, R.F., Afzal, V., Sadinski, W., Trask, B.J., Legerski, R. and Cleaver, J.E.
TITLE Complementation of chromosomal instability in the xeroderma pigmentosum variant by a gene on human chromosome 1 with homology to S-adenosyl homocysteine hydrolase
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 2258)
AUTHORS Volpe, J.P.G., McDowell, M. and Cleaver, J.E.
TITLE Direct Submission
JOURNAL Submitted (19-DEC-1996) Dermatology, UCSF, 3rd and Parnassus, Box 0750, San Francisco, CA 94143, USA
FEATURES
source
1. .2258
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/cell_type="fibroblasts"
/tissue_type="skin"
1. .2258
/gene="XPVKona"
48. .1550
/gene="XPVKona"
/note="xeroderma pigmentosum variant"
/codon_start=1
/product="S-adenosyl homocysteine hydrolase homolog"
/protein_id="AAC01960.1"
/db_xref="GI:2852125"
/translation="MATVTKAPKKQIQFADDMQEFTEPKTGRSLRSISQSSDLS
YSSAASYTDSDEVSPEKQNTSKSSNFCVKNIKQAEGRREIEIAEQDMSALIS
LKRRAQGEKPLAGAKIVGCTHTTAQTAVLIETLICALGACRWSACNISTONEVAAL
AEAGVAVFAWKGESEDDFWCIDRCVNMDSVTKOKFDNLCCRESILDGLKRTTD
RGIVESVTGVHRLYQLSKAGKLCVPMANVNDSTKOKFDNLCCRESILDGLKRTTD
VMFGKQVVCVCGYGEVGGCCALKAIGAIYVITEIDPICALQACMDGFRVVKLNEYI
ROVDVITCTGNKNVTRREHLDKMSKSCIVCMGHSNTEIDVTSRLTPELTWERVRSQ
VDHVIMPDKRVLVLAEGRLNLSCSTVPTFVLSITATQALALIELYNAPEGRYKQD
VYLLPKMDEYVASLHLPFDALHLELTDQAKYLGINKNGPFKNYYRY"

BASE COUNT 603 a 513 c 542 g 600 t
ORIGIN

Query Match 100.0%; Score 417; DB 9; Length 2258;
Best Local Similarity 100.0%; Pred. No. 3.7e-117;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAGAAGCAGCAACCACTCCAGAGGAGCAGCAGCAATTTCTGTGGAAGACATCAAGC 60
DB 232 GAGAGAAGCAGCAACCACTCCAGAGGAGCAGCAGCAATTTCTGTGGAAGACATCAAGC 291
QY 61 AGGCAGAAATTGGACGCCGGGAGATTGAGATTGCAGAGCAAGACATGCTCTGATTT 120
DB 292 AGGCAGAAATTGGACGCCGGGAGATTGAGATTGCAGAGCAAGACATGCTCTGATTT 351
QY 121 CACTCAGGAAACGTGCTCAGGGGAGAAAGCCCTGGCTGGTCTAAATAAGTGGGCTGTA 180
DB 352 CACTCAGGAAACGTGCTCAGGGGAGAAAGCCCTGGCTGGTCTAAATAAGTGGGCTGTA 411
QY 181 CACACATCACAGCCAGACAGCGGTGTTGATTGAGACACTCTGCCCCCTGGGGCTCAGT 240
DB 412 CACACATCACAGCCAGACAGCGGTGTTGATTGAGACACTCTGCCCCCTGGGGCTCAGT 471
QY 241 GCCGCTGCTGCTGTGAACATCTACTCACTCAGAAATGAAGTAGCTGCAGCAGCTGGCTG 300
DB 472 GCCGCTGCTGCTGTGAACATCTACTCACTCAGAAATGAAGTAGCTGCAGCAGCTGGCTG 531
QY 301 AGGCTGGAAGTTCAGAGTGTCCGTTGGAAGGGGAGTCAAGATGACTTCTGTGTGTGTA 360
DB 532 AGGCTGGAAGTTCAGAGTGTCCGTTGGAAGGGGAGTCAAGATGACTTCTGTGTGTGTA 591

QY 361 TTGACCGCTGTGTGAACATGATGGGTGGCAGGCCCAACATGATCCTGATGATGGGG 417
DB 592 TTGACCGCTGTGTGAACATGATGGGTGGCAGGCCCAACATGATCCTGATGATGGGG 648

RESULT 2
BC007576
LOCUS
DEFINITION
Homo sapiens, S-adenosylhomocysteine hydrolase-like 1, clone
MGC:15558 IMAGE:3139729, mRNA, complete cds.
ACCESSION
BC007576
VERSION
BC007576.1 GI:14043176
KEYWORDS
MGC.
SOURCE
Homo sapiens.
ORGANISM
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo;
1 (bases 1 to 2526)
Strausberg, R.
Direct Submission
Submitted (10-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgapsb-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Heiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeary, Steven
Ness, Pawan Pandoh, Anna-Lissa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smalus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Matsya van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

REMARK
COMMENT

FEATURES
source
1. .2526
/organism="Homo sapiens"
/db_xref="locusID:10768"
/db_xref="taxon:9606"
/clone="MGC:15558 IMAGE:3139729"
/tissue_type="Placenta, chorlocarcinoma"
/clone_lib="NIH_MGC_21"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
308. .1810
/codon_start=1
/product="S-adenosylhomocysteine hydrolase-like 1"
/protein_id="AAH07576.1"
/db_xref="GI:14043177"
/translation="MATVTKAPKKQIQFADDMQEFTEPKTGRSLRSISQSSDLS
YSSAASYTDSDEVSPEKQNTSKSSNFCVKNIKQAEGRREIEIAEQDMSALIS
LKRRAQGEKPLAGAKIVGCTHTTAQTAVLIETLICALGACRWSACNISTONEVAAL
AEAGVAVFAWKGESEDDFWCIDRCVNMDSVTKOKFDNLCCRESILDGLKRTTD
RGIVESVTGVHRLYQLSKAGKLCVPMANVNDSTKOKFDNLCCRESILDGLKRTTD
VMFGKQVVCVCGYGEVGGCCALKAIGAIYVITEIDPICALQACMDGFRVVKLNEYI
ROVDVITCTGNKNVTRREHLDKMSKSCIVCMGHSNTEIDVTSRLTPELTWERVRSQ
VDHVIMPDKRVLVLAEGRLNLSCSTVPTFVLSITATQALALIELYNAPEGRYKQD
VYLLPKMDEYVASLHLPFDALHLELTDQAKYLGINKNGPFKNYYRY"

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>
Series: IRAL Plate: 22 Row: h Column: 21
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, similarity but not identity to protein.

BASE COUNT 644 a 592 c 660 g 630 t
ORIGIN

Query Match 100.0%; Score 417; DB 9; Length 2526;
Best Local Similarity 100.0%; Pred. No. 3.7e-117;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAGAAGCAGCAACCACTCCAGAGGAGCAGCAATTTCTGTGAAGAATCAAGC 60
DB 492 GAGAGAAGCAGCAACCACTCCAGAGGAGCAGCAATTTCTGTGAAGAATCAAGC 551
QY 61 AGCAGAATTTGAGCGCGGAGATGAGATTGACAGCAAGACATGCTGCTGATTT 120
DB 552 AGCAGAATTTGAGCGCGGAGATGAGATTGACAGCAAGACATGCTGCTGATTT 611
QY 121 CACTCAGGAACGCTGCTCAGGGGAGAGACCCCTGGCTGCTAAATAAGTGGCTGTA 180
DB 612 CACTCAGGAACGCTGCTCAGGGGAGAGACCCCTGGCTGCTAAATAAGTGGCTGTA 671
QY 181 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTCTGCTGCTGGGGCTCAGT 240
DB 672 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTCTGCTGCTGGGGCTCAGT 731
QY 241 GCCCGTGTGCTGTTGAACATCTACTCAACTCAGAAATGAGTACAGCAGCAGTGGCTG 300
DB 732 GCCCGTGTGCTGTTGAACATCTACTCAACTCAGAAATGAGTACAGCAGCAGTGGCTG 791
QY 301 AGCGTGAGTTGAGTGTGCTTGAAGGCGAGTCAAGAGATGACTTCTGTGTGTA 360
DB 792 AGCGTGAGTTGAGTGTGCTTGAAGGCGAGTCAAGAGATGACTTCTGTGTGTA 851
QY 361 TTGACCGCTGTGTGAACATGATGGTGGCAGGCCAATGATCTGTGATGGGG 417
DB 852 TTGACCGCTGTGTGAACATGATGGTGGCAGGCCAATGATCTGTGATGGGG 908

RESULT 3
LOCUS BC016942 2552 bp mRNA linear PRI 09-NOV-2001
DEFINITION Homo sapiens, S-adenosylhomocysteine hydrolase-like 1, clone
MGC:21453 IMAGE:3450568, mRNA, complete cds.
ACCESSION BC016942
VERSION BC016942.1 GI:16877386
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 2552)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (05-NOV-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT NTH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgaphs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcdpaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAK Plate: 20 Row: a Column: 15
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not

FEATURES
Source Identity to protein.
Location/Qualifiers
1..2552

/organism="Homo sapiens"
/db_xref="LocusID:10768"
/db_xref="taxon:9606"
/clone="MGC:21453 IMAGE:3450568"
/tissue_type="Placenta, choriochorionoma"
/clone_1lb="NTH MGC_10"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
338..1840
/codon_start=1
/product="S-adenosylhomocysteine hydrolase-like 1"
/protein_id="AAH16942.1"
/db_xref="GI:16877387"

CDS

/translation="MATVTKAPKKQIQFADDMOEFTRKTRRSLSRSISQSDTS
YSSASVTDSSDDEVSPREKQTNKSSNFCVKNIKQAEGRREIEAEDMSALIS
LRRKQGEKPLAGAKIVGCTHITAGTAVLIELCALGAOCRSACNISTONEVAAL
AEAGVAFAMKGESEDEDFWNCIDRCVNDGQANMILDDGDLTHWVYKKYPNFKKI
RGIVESVTGVHRLYQLSKAGKLYPAMNVDVTKQKFDNLCCRESLDGLKRTD
VMEGKQVAVVCGYGEVKGKCCALALGAIIVITEIDPICALQACMDGFRVVKLEVI
ROYVYITCTGNKNVTRHELDKNSCIYCNMGSNTEIDYSLRTPELTWERVRSQ
VDHVIWPDGKRVVLLAEGRLNLSCTVPTFVLSITATLQALALELYNAPEGRYKOD
VYLLPKMDEVASLHLPSEDAHLELTDQAKYLGINKNGPERPNYRY"

BASE COUNT 644 a 600 c 671 g 637 t
ORIGIN

Query Match 100.0%; Score 417; DB 9; Length 2552;
Best Local Similarity 100.0%; Pred. No. 3.7e-117;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAGAAGCAGCAACCACTCCAGAGGAGCAGCAATTTCTGTGAAGAATCAAGC 60
DB 522 GAGAGAAGCAGCAACCACTCCAGAGGAGCAGCAATTTCTGTGAAGAATCAAGC 581
QY 61 AGCAGAATTTGAGCGCGGAGATGAGATTGACAGCAAGACATGCTGCTGATTT 120
DB 582 AGCAGAATTTGAGCGCGGAGATGAGATTGACAGCAAGACATGCTGCTGATTT 641
QY 121 CACTCAGGAACGCTGCTCAGGGGAGAGAGCCCTGGCTGCTAAATAAGTGGCTGTA 180
DB 642 CACTCAGGAACGCTGCTCAGGGGAGAGAGCCCTGGCTGCTAAATAAGTGGCTGTA 701
QY 181 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTCTGCTGCTGGGGCTCAGT 240
DB 702 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTCTGCTGCTGGGGCTCAGT 761
QY 241 GCCCGTGTGCTGTTGAACATCTACTCAACTCAGAAATGAGTACAGCAGCAGTGGCTG 300
DB 762 GCCCGTGTGCTGTTGAACATCTACTCAACTCAGAAATGAGTACAGCAGCAGTGGCTG 821
QY 301 AGCGTGAGTTGAGTGTGCTTGAAGGCGAGTCAAGAGATGACTTCTGTGTGTA 360
DB 822 AGCGTGAGTTGAGTGTGCTTGAAGGCGAGTCAAGAGATGACTTCTGTGTGTA 881
QY 361 TTGACCGCTGTGTGAACATGATGGTGGCAGGCCAATGATCTGTGATGGGG 417
DB 882 TTGACCGCTGTGTGAACATGATGGTGGCAGGCCAATGATCTGTGATGGGG 938

RESULT 4
LOCUS AX029176 2563 bp DNA linear PAT 16-SEP-2000
DEFINITION Sequence 1 from Patent WO9814562.
ACCESSION AX029176
VERSION AX029176.1 GI:10190060
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 2563)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS	Hart, D. N.
TITLE	Enzyme having s-adenosyl-l-homocysteine hydrolase (ahcy) type activity
JOURNAL	Patent: WO 9814562-A 1 09-APR-1998;
FEATURES	HART, DEREK NIGEL JOHN (NZ)
source	Location/Qualifiers 1. .2563

CDS

/note="Open reading frame extends without a stop codon for the full 5' nucleotide sequence. The initiation codon has yet to be identified."

```

/protein_id="CAC09285.1"
/db_xref="GI:10190061"
/translation="RGQVGARSCCFWFSCGHRCPALGCRITDKAMAPLAPQKPTQLDA
GAGRRVGRVSEGAARGRAPEGEWGGGGGSAAGRAGGMSMPDAMPAPLPGVEELKQ
AKIEDAEKYSFMATVTKAPKKQIQFADDMQEFTEKPTKIGRRSLSNSISQSTDSYS
SAASYTDSDEVSPREKQQTNSGNSFCVKNIKQAEFGREIEIEMQDSALISLR
KRAOGEKPLAGAKIVGCTHTIQAVALIETLCALGACQWMSACNYSITONEVAALAE
AGAAVEAMKGESEDEFWMCIDRCVNMNDVQANMILDDGGDLTHVYKKYPNVEFKKIR
IVEESVTGVHRLYQLSKRAGKLCVPAMVNDSVTKQKEDNLCCRESLIDGLKRTTDMV
GGKQGVVCGYGEVGKCCALAKALGAIVYITEIDPICALQACMDGFVYKLVNIYRQ
VDVATITGKNKVLAEGRHLNDRKNSCIIVCMHSGATSTQDALVSLRTPELTWERVRQVD
HYIWPDGKRVLLAEGRHLNDRKNSCIIVTFLVLSITATQDALVIELNAPEGRYKQDYY
LIPKKDEIVASLHLPSEDAHLTEIJDQAKYLGINKNGPEKPNVYR"

```

BASE COUNT	646 a	604 c	677 g	636 t
ORIGIN				

Query Match	100.0%;	Score 417;	DB 6;	Length 2563;
Best Local Similarity	100.0%;	Pred. No. 3.7e-117;		
Matches 417;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

OY	1	GAGAGAAGCAGCAACCACTCCAAAGGGCAGCAGCAATTCTGTGTGAAGAACATCAAGC	60
Db	529	GAGAGAAGCAGCAACCACTCCAAGGGCAGCAGCAATTCTGTGTGAAGAACATCAAGC	588
OY	61	AGGCAGAAITTTGGACGCCCCGGAGATTGAGATTGCAGAGCAAGACATGTCTGCTTGATTT	120
Dd	589	AGGCAGAAITTTGGACGCCCGGGAGATTGAGATTGCAGAGCAAGACATGTCTGCTTGATTT	648
OY	121	CACCTCAGGAACGTGCTCAGGGGGAGAACCCCTTGGCTGTGCTTAATAATAGTGGGCTGTA	180

Db	649	CAC	TCA	GGA	AA	CG	TG	CT	CAG	GGG	GA	AA	GC	CT	TG	GC	TAA	AA	TAG	TGG	GG	CT	GTA	708			
OY	181	CAC	ACAT	CAC	AG	CC	CAG	AC	AG	CG	GT	GT	GAT	TG	AG	AC	ACT	CT	GT	GC	CC	CT	GG	GG	CT	CAGT	240
Db	709	CAC	ACAT	CAC	AG	CC	CAG	AC	AG	CG	GT	GT	GAT	TG	AG	AC	ACT	CT	GT	GC	CC	CT	GG	GG	CT	CAGT	768

QY	241	6CCGCTGCTGCTGTAACTACTCAGATGAAGTAGCTGCAGCACTGGCTG	300
Db	769	6CCGCTGCTGCTGTAACTACTCAGATGAAGTAGCTGCAGCACTGGCTG	828

QY	301	AGCGTGGAGTTGCAGTGTTCCTTGGAAGGGCGAGTCAGAGATGACTCTGTGCTGTGA	360
DB	829	AGCGTGGAGTTGCAGTGTTCCTTGGAAGGGCGAGTCAGAGATGACTCTGTGCTGTGA	888

Qy	361	TTGACCGCTGTGTGAACATGGATGGGTGGCAGCCAAACATGATCCTGGATGATGGGG	417
Db	889	TTGACCGCTGTGTGAACATGGATGGGTGGCAGCCAAACATGATCCTGGATGATGGGG	945

RESULT 5
BC010681

LOCUS	BC010681	2583 bp	mRNA	linear	PRI 12-JUL-2001
DEFINITION	Homo sapiens, S-adenosylhomocysteine hydrolase-like 1, clone MGC:8936 IMAGE:3853747, mRNA, complete cds.				

ACCESSION BC010681
VERSION BC010681.1 GI:14715037
KEYWORDS
SOURCE MGC.
ORGANISM Homo sapiens.
Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2583)
Strausberg, R.
Direct Submission
Submitted (10-JUL-2001) National Institutes of Health, Mammalian

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>

Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: villalonaebcm.tmc.edu.
Villalón, D.K., Luna, B.A., Hale, S.M., Hu]

A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny D.M., Gibbs P.A.

klondike/klondike: MC

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.jnl.gov>
Series: IRAK Plate: 13 Row: b Column: 21
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA 91: 5729723.

FEATURES

Source

```

source
1. .2583
/organism="Homo sapiens"
/db_xref="LocusID:10768"
/db_xref="taxon:9606"
/clone="MGC:8936 IMAGE:3853747"
/tissue_type="Colon, adenocarcinoma"
/clone_1lb="NH_MGC_65"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
367. .1869
CDS

```

CDS

BASE COUNT	648 a	606 c	688 g	641 t
ORIGIN				

Query Match	100.0%;	Score 417;	DB 9;	Length 2583;
Best Local Similarity	100.0%;	Pred. No. 3.7e-117;		
Matches 417; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0

QY 1 GAGAGAAGCAGCAACCACTCCAAAGGGCAGCAGCAATTTCTGTGAGAACAATCAAGC 60
|||||
Db 551 GAGAGAAGCAGCAACCACTCCAAAGGGCAGCAGCAATTTCTGTGTGAGAACAATCAAGC 610

[illegible][illegible]

QY 181 CACACATCACAGCCCGGTGTGATTGAGACACTCTGTGCCCTGGGGCTCAGT 2400

```
|||||
Db 731 CACACATCACAGCCAGACAGCGGCTGTGATGAGACACTCTGTGCCCTGGGGCTCAGT 790
QY 241 GCCCGTGGTCTGCTGTGTACATCTTACTCACTCAGATGAGTAGCTGACGACTGGCTG 300
Db 791 GCCCGTGGTCTGCTGTGTACATCTTACTCACTCAGATGAGTAGCTGACGACTGGCTG 850
QY 301 AGCGTGGAGTGTGAGTGTGCTGGAGAGGGCGAGTCAAGATGACTTCTGTGTGTGA 360
Db 851 AGCGTGGAGTGTGAGTGTGCTGGAGAGGGCGAGTCAAGATGACTTCTGTGTGTGA 910
QY 361 TTGACCGCTGTGTGACATGATGGGTGGCAGCCACATGATCCTGTGATGATGGGG 417
Db 911 TTGACCGCTGTGTGACATGATGGGTGGCAGCCACATGATCCTGTGATGATGGGG 967

RESULT 6
AF315687 2677 bp mRNA linear PRI 22-MAR-2002
LOCUS Homo sapiens S-adenosylhomocysteine hydrolase-like protein mRNA,
DEFINITION complete cds.
ACCESSION AF315687
VERSION AF315687.1 GI:16588686
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2677)
AUTHORS Dekker,J.W., Budhia,S., Angel,N.Z., Cooper,B.J., Clark,G.J.,
Hart,D.N. and Kato,M.
TITLE Identification of an S-adenosylhomocysteine hydrolase-like
transcript induced during dendritic cell differentiation
JOURNAL Immunogenetics 53 (12), 993-1001 (2002)
MEDLINE 21901265
PUBMED 11904675
REFERENCE 2 (bases 1 to 2677)
AUTHORS Dekker,J.W., Budhia,S., Angel,N.Z., Cooper,B.J., Clark,G.J.,
Hart,D.N.J. and Kato,M.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-2000) Dendritic Cell Research, Mater Medical
Research Institute, Level 3, Aubigny Place, South Brisbane,
Queensland 4101, Australia

FEATURES
Source
1..2677
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p21"
/clone="211(1)B"
/cell_line="Hodgkin's disease cell line LA28"
1..114
/note="derived from 5' RACE"
369..1961
/note="DCAL, dendritic cell expressed AHCT-like protein;
expressed in dendritic cells in blood, Langerhans cells
and dermal dendritic cells; mRNA increased during monocyte
differentiation to monocyte-derived dendritic cells in the
presence of GM-CSF and IL-4"
/codon_start=1
/product="S-adenosylhomocysteine hydrolase-like protein"
/protein_id="AAL26869.1"
/db_xref="GI:16588687"
/translation="MSMPDAMPLEPGVEELKQAEIEDAEKISFNAVTKAPKKQIOF
ADDMQETKPTKTRSLRSISGSDSYSSASITDSSDEVSPPREKQNTSKGS
SNFCVKNIKQAEFGREIEIAEDMSALISLRKRAQEKPLAGAKIVGCTHTAQTAV
LLETICALGAQCRWSACNITSTQNEVAAALAEAGVAVFAWKGESEDFWMCIDRCVM
DGMQANMILDDGDLTHWVYKYPNFKIRGIVESVTGVHLYOLSKAGKLCVPAM
NVNDSVTKQKFDNLVCCRESILDLKRTTDMFGKQVVCVGGYGVKGCACALKALG
AIVYTEIDPICALQACMDGFRVYKNEVIRQVDVITCTGNKNVYTRHLDRMKNSC
IVCNMGHSNTEIDVTSRLTPELTWERVSQVDHVIWPDGKRVVLLAEGRLNLSCSTV
PEFVLSITATQALALIELYNAPGRKQDVYLLPKMDEVYASLHLPFDAHLTELT
DDQAKYLGINKNGPFKNYYRY"
```

```
BASE COUNT 657 a 648 c 708 g 664 t
ORIGIN
Query Match 100.0%; Score 417; DB 9; Length 2677;
Best Local Similarly 100.0%; Pred. No. 3.7e-117;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAGAAGCAGCAACCACTCCAAAGGGCAGCAGCAATTCTGTGTGAAGAATCAAGC 60
Db 643 GAGAGAAGCAGCAACCACTCCAAAGGGCAGCAGCAATTCTGTGTGAAGAATCAAGC 702
QY 61 AGGAGAATTGGACGGCGGAGATTGAGATTGCAGACAAGACATGCTGCTCTGATTT 120
Db 703 AGGAGAATTGGACGGCGGAGATTGAGATTGCAGACAAGACATGCTGCTCTGATTT 762
QY 121 CACTCAGAAACGTCGTCAGGGGAGAGAGCCCTTGCTGCTAAATAGTGGCTGTA 180
Db 763 CACTCAGAAACGTCGTCAGGGGAGAGAGCCCTTGCTGCTAAATAGTGGCTGTA 822
QY 181 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGGGGCTCAGT 240
Db 823 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGGGGCTCAGT 882
QY 241 GCCCGTGGTCTGCTGTGTACATCTTACTCACTCAGATGAGTAGCTGCAGACTGGCTG 300
Db 883 GCCCGTGGTCTGCTGTGTACATCTTACTCACTCAGATGAGTAGCTGCAGACTGGCTG 942
QY 301 AGCGTGGAGTGTGAGTGTGCTGGAGAGGGCGAGTCAAGATGACTTCTGTGTGTGA 360
Db 943 AGCGTGGAGTGTGAGTGTGCTGGAGAGGGCGAGTCAAGATGACTTCTGTGTGTGA 1002
QY 361 TTGACCGCTGTGTGACATGATGGGTGGCAGCCACATGATCCTGTGATGATGGGG 417
Db 1003 TTGACCGCTGTGTGACATGATGGGTGGCAGCCACATGATCCTGTGATGATGGGG 1059

RESULT 7
HSM800298 2510 bp mRNA linear PRI 18-FEB-2000
LOCUS Homo sapiens mRNA; cDNA DKFZp564A1523 (from clone DKFZp564A1523);
DEFINITION partial cds.
ACCESSION AL049954
VERSION AL049954.1 GI:4884203
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2510)
AUTHORS Wambutt,R., Heubner,D., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
TITLE Direct Submission
JOURNAL Submitted (15-MAY-1999) MIPS, Am Klopferplatz 18a, D-82152
Martinsried, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
This clone (DKFZp564A1523) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://www.mips.biochem.mpg.de/proj/cDNA/.

FEATURES
Source
1..2510
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="1, 15.55 cR from DIS248"
/clone="DKFZp564A1523"
/tissue="brain"
/clone_lib="564 (synonym: hfbz2). Vector PAMP1; host
X1-2blue; sites NotI + SalI"
/dev_stage="fetal"
1..1796
gene
```

CDs

/gene="DKFZp564A1523"
<1.1796
/gene="DKFZp564A1523"
/note="S-adenosyl homocysteine hydrolase homolog XPVKona
splice variant"
/codon_start=3
/product="hypothetical protein"
/protein_id="CAB43223.1"
/db_xref="GI:4884204"
/translation="GLTVELDSCVYILLRLHRGSRGKICAEISQAPLPQE
TQAPSGESENRLRLRQGRRAESSNLISRLVSAQNGRSRWMLIQLPEIDSL
ISCFIQIADDMQEFTEPTKTRSLRSISQSDSYSSASTDSDDEVSPRE
KQOTNSKGSNFCVKNIKQAEFGREIEIAEODMSALISLRKAGEKPLAGAKIVC
THTRQTAVLIELTLCALGAQCRSACNISTONEVAALAEAGVAFWKGESEDFW
WCIDRCVMDGQANMILDDGDLTHWYKKYPNVEFKIRGIVEEYVGHRLYOLSK
AGKLCVPMNVNDSVTKQKFDNLCCRESILDLKRTDVMFGKQVYVCGYGEVKG
CCAAIKALGAIVYTEIDPICALQACMDGFRVVKLNEVIRQVDVITCTGNKNVTR
HLDRMNSCIVCMGHSNTEIDVTSLTPELTWERVRSQVDHVIWPDGKRVLLAEG
LNLSCSTVPTFLSITATQALILELYNAPEGRYKQDVYLLPKMDEVVASLHLPS
FDAHLELTDDQAKYLGINKNGPFPNYYRY"

polyA_site 664 a 556 c 618 g 672 t

BASE COUNT 664 a 556 c 618 g 672 t

ORIGIN

Query Match 99.6%; Score 415.4; DB 9; Length 2510;
Best Local Similarity 99.8%; Pred. No. 1.2e-116;
Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGAGAGCAGCAACCACTCCAAAGGCGACGACCAATTCTGTGTGAGAACATCAAGC 60
|||||
DB 478 GAGAGAGCAGCAACCACTCCAAAGGCGACGACCAATTCTGTGTGAGAACATCAAGC 537

QY 61 AGGCAGATTGTGACGGCGGAGATTGAGATTGACAGACAGACATGCTGCTGTGATT 120
|||||
DB 538 AGGCAGATTGTGACGGCGGAGATTGAGATTGACAGACAGACATGCTGCTGTGATT 597

QY 121 CACTCAGGAACGCTCAGGGGAGAGAGCCCTGCTGCTAAATAGTGGGCTGTA 180
|||||
DB 598 CACTCAGGAACGCTCAGGGGAGAGAGCCCTGCTGCTAAATAGTGGGCTGTA 657

QY 181 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGGGGCTCAGT 240
|||||
DB 658 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGGGGCTCAGT 717

QY 241 GCCGCTGCTGCTGTTGTAACATCTACTCACTCAGAAATGAGTAGCTGCAGCAGCTGGCTG 300
|||||
DB 718 GCCGCTGCTGCTGTTGTAACATCTACTCACTCAGAAATGAGTAGCTGCAGCAGCTGGCTG 777

QY 301 AGGCTGAGTTGACGTGCTGTTGAGAGGGCGAGTCAGAGATGACTTCTGGTGTGTA 360
|||||
DB 778 AGGCTGAGTTGACGTGCTGTTGAGAGGGCGAGTCAGAGATGACTTCTGGTGTGTA 837

QY 361 TTGACCGCTGTGTGACATGATGGGTGCGAGGCCAACATGATCCTGATGATGGG 417
|||||
DB 838 TTGACCGCTGTGTGACATGATGGGTGCGAGGCCAACATGATCCTGATGATGGG 894

RESULT 8
BC018218 1772 bp mRNA linear ROD 07-AUG-2002
LOCUS Mus musculus, S-adenosylhomocysteine hydrolase-like 1, clone
DEFINITION MGC:18748 IMAGE:4007102, mRNA, complete cds.
ACCESSION BC018218
VERSION BC018218.1 GI:17390492
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1772)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (03-DEC-2001) National Institutes of Health, Mammalian

REMARK
COMMENT

Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIR-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunnarone, P.H., Garcia, A.M., Lu, X., Huylk, S.W., Hale, S.M.,
Yoon, V.S., Kowls, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>.
Series: IRAK Plate: 23 Row: n Column: 6
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Genomescan gene prediction, similarity but not identity
to protein.

FEATURES
source

location/Qualifiers
1.1772
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="CZECH II"
/clone="MGC:18748 IMAGE:4007102"
/tissue_type="Mammary tumor metastasized to lung."
MTV-LTR/Whit model. Expression driven by an MTV-LTR
enhancer."
/clone_1lb="NCI CGAP_Lu30"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
191.1693
/codon_start=1
/product="S-adenosylhomocysteine hydrolase-like 1"
/protein_id="AAH18218.1"
/db_xref="GI:17390493"
/db_xref="LocusID:10768"
/translation="MATVTKAPKKOIOFADDMQEFTEPTKTRSLRSISQSDTS
YSSASTDSDDEVSPREKQOTNSKGSNFCVKNIKQAEFGREIEIAEODMSALIS
LRKRAOGEKPLAGAKIVCTHITAOAVLIELTLCALGAQCRSACNISTONEVAAL
AEAGVAVFAWKGESEDFWICIDRCVMDGQANMILDDGDLTHWYKKYPNVEKRI
RGIVEEYVGHRLYOLSKAGKLCVPMNVNDSVTKQKFDNLCCRESILDLKRTD
VMFGKQVYVCGYGEVKGCCAAIKALGAIVYTEIDPICALQACMDGFRVVKLNEVI
ROVDVITCTGNKNVTRHLDRLMNSCIVCMGHSNTEIDVTSLTPELTWERVRSQ
VDHVIWPDGKRVVLLAEGRLNLNLSCSTVPTFLSITATQALILELYNAPEGRYKOD
VYLLPKMDEVVASLHLPSFDAHLELTDDQAKYLGINKNGPFPNYYRY"

CDS

BASE COUNT 462 a 406 c 506 g 398 t

ORIGIN

Query Match 92.3%; Score 385; DB 10; Length 1772;
Best Local Similarity 95.2%; Pred. No. 2.6e-107;
Matches 397; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 GAGAGAGCAGCAACCACTCCAAAGGCGACGACGACCAATTCTGTGTGAGAACATCAAGC 60
|||||
DB 375 GAGAGAGCAGCAACCACTCCAAAGGCGACGACGACCAATTCTGTGTGAGAACATCAAGC 434

QY 61 AGGCAGATTGTGACGGCGGAGATTGAGATTGACAGACAGACATGCTGCTGATT 120
|||||
DB 435 AGGCAGATTGTGACGGCGGAGATTGAGATTGACAGACAGACATGCTGCTGATT 494

QY 121 CACTCAGGAACGCTCAGGGGAGAGAGCCCTGCTGCTAAATAGTGGGCTGTA 180
|||||
DB 495 CACTCAGGAACGCTCAGGGGAGAGAGCCCTTGGCTGCTAAATAGTGGGCTGTA 554

QY 181 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGGGGCTCAGT 240
|||||

Db 555 CACACATCACGGCCAGACAGCGGATTAATTGAGACCCCTTGTGCCCTGGGAGCTCAGT 614
QY 241 GCCGTGGTCTGCTGTACATCTCACTCAGATGAAGTAGTGCAGCACTGGCTG 300
Db 615 GCCGTGGTCTGCTGCAACATCTATCACTCAGATGAAGTAGTGCAGCACTGGCTG 674
QY 301 AGGCTGAGTTGACGTCTGCTGGAAGGGCGAGTCAGAGATGACTCTGCTGCTGTA 360
Db 675 AGGCTGAGTCCGCGTGTGCTGGAAGGGCGAGTCAGAGATGACTCTGCTGCTGCA 734
QY 361 TTGACCGCTGTGTGACATGATGGTGGCAGGCCAATGATCTCTGATGATGGG 417
Db 735 TTGACCGCTGTGTGACATGATGGTGGCAGGCCAATGATCTCTGATGATGGG 791

RESULT 9
AK025372
LOCUS AK025372 1987 bp mRNA linear PRI 29-SEP-2000
DEFINITION Homo sapiens CDNA: FLJ21719 fls, clone COLF0094.
ACCESSION AK025372
VERSION AK025372.1 GI:10437875
KEYWORDS oligo capping; fls (full insert sequence).
SOURCE Homo sapiens colon mucosa CDNA to mRNA, clone_1lb:COLF
clone:COLF0094.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (sites)
Watanabe, K., Kumagai, A., Itakura, S., Yamazaki, M., Tashiro, H.,
Ota, T., Suzuki, Y., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,
Nakamura, Y., Isogai, T. and Sugano, S.
NEDO human CDNA sequencing project
TITLE Unpublished
JOURNAL 2 (bases 1 to 1987)
REFERENCE Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isogai, T.,
AUTHORS Shibahara, T., Tanaka, T. and Nakamura, Y.
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail: cdna@genome.u-tokyo.ac.jp, Tel: 81-3-5449-5286,
Fax: 81-3-5449-5416)
COMMENT NEDO human CDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; CDNA full insert
sequencing; Research Association for Biotechnology; CDNA library
construction, 5'- & 3'-end one pass sequencing; Departent of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).

FEATURES
Source Location/Qualifiers
1. 1987
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="COLF0094"
/tissue_type="colon mucosa"
/clone_1lb="COLF"
/note="cloning vector pME18SFL3"
BASE COUNT 557 a 430 c 501 g 499 t
ORIGIN

Query Match: 66.2%; Score 276.2; DB 9; Length 1987;
Best Local Similarity 78.9%; Pred. No. 8e-74;
Matches 329; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 1 GAGAGAGCAGCAAAACCACTCCAGGGCAGCAGCAATTTCTGTGAAGACATCAAGC 60
Db 480 GGGACAGCAGCAAAAGAACTCTAAGGGAGCAGTGACTTCTGTGAAGACATCAAAC 539
QY 61 AGGCAGAATTTGACGCCGGAGATTGAGATTGCAGAGCAAGACATGCTCTGATTT 120
Db 540 AGGCAGAATTTGACAGCAAGAAATTTGCTGAACAAGAAATGCTGATGATGG 599
QY 121 CACTCAGGAACGCTGCTCAGGGGGAGAACCCCTTGCTGTCTAAATAAGTGGGCTGTA 180

Db 600 CTTGAGGAGAGAGCTCAAGAGAAAAGCCCTTGGCTGAGCCAAAATCGTGGTTGCA 659
QY 181 CACACATCACAGCCAGACAGCGGCTGTGATTGAGACACTCTGTGCCCTGGGGCTCAGT 240
Db 660 CACACATCACCTGCAAGACTGCTGTGTATGGAACCTGCGTCTGGGGCCAGT 719
QY 241 GCCGTGGTCTGCTGTACATCTCACTCAGATGAAGTAGTGCAGCACTGGCTG 300
Db 720 GCCGATGGCTGCTGCAACATCTATTCCACTCTCAATGAAGTGGCTGCTGCTAGCAG 779
QY 301 AGGCTGAGTTGACGTCTGCTGGAAGGGCGAGTCAGAGATGACTCTGCTGCTGTA 360
Db 780 AAAGTGATTTCTGTTTGGCTTGAAGGAGAGAGATGACTCTGCTGCTGTA 839
QY 361 TTGACCGCTGTGTGACATGATGGTGGCAGGCCAATGATCTCTGATGATGGG 417
Db 840 TCGATAGATGTGTGAATGTGAGAGGCGTGGCAGCCAAACATGATCTTGATGATGAG 896

RESULT 10
BC008349
LOCUS BC008349 2030 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, similar to S-adenosylhomocysteine hydrolase-like 1,
clone IMAGE:3536052, mRNA, partial cds.
ACCESSION BC008349
VERSION BC008349.1 GI:14249935
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2030)
Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (25-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guln,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

FEATURES
source Location/Qualifiers
1. 2030
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3536052"
/tissue_type="Lung, small cell carcinoma"
/clone_1lb="NIH_MGC_7"
/lab_host="DH10B-R"
/note="Vector: pOT87"
/codon_start=3
/product="Similar to S-adenosylhomocysteine hydrolase-like
1"

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 23 Row: J Column: 18.

```

/protein_id="AAH08349.1"
/db_xref="GI:14249936"
/translation="MSVQVPSAAAAAKVEVELKDLSPSEAESQGLSTAAGCAMAAPP
AGGDPPEAPAPAERPPVQVPGSGPAALSPAAGKVPQASAMKRSDPHHQHRDRDGG
EALVSPDGTVEAPRTVKKQIQFADQKEFNKPTKIGRSLRSISQSTDSYSSAA
SYTDSDEDTSPRDKQOKNSKSSDFCYKNIKQAEFGRREIEIAEQEMALMLRKRA
OGEKPLAGAKIVGCTHITAOATAVLMETGALGAQCRAWACNIYSTLINEVAALAESGF
PVEAKGESEDDFWMCIDRCVANEVGMQPMNLLDDGDLTHWIKKYPNPKIKIGIVE
ESVTGVHRLYQLSKAGKLCVPAWNVYDSTKQKEDNLCCRESILDGKRTTDMFEGE
KQVVVCGYGEVGGCCAAALKAMGSIYVTEIDPICALQACMGSEFLVGLKRTIDMEDI
VITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVASLRTPELTWERVRSOVDHVI
WPDGKRIYLLAEGRLNLSCSTVPTFVLSITATQALALIELYNAPEGRRKQDYYLLP
KKMDEYASLHLPTFDAHLELTDEQAQYGLNKNGPFPKNYYRY"

```

Query Match	65.98;	Score 274.6;	DB 9;	Length 2030;
Best Local Similarity	78.78;	Pred. No. 2.5e-73;		
Matches 328;	Conservative	0;	Mismatches 89;	Indels 0;
				Gaps 0;

QY	1	GAGAGAAGCAGCAAAACCAACTCCAAAGGGCAGCAGCAATTTCTGTGTGAAGAATCAAGC	60
Db	520	GGGACAAAGCAGCAAAAGAACTCTAAGGGAAAGCAGTGACTTCTGTGTAAAGACATCAAAAC	579
QY	61	AGGCAGAAATTTGGACGCGCGGAGATTGAGATTGCAGAGCAAGACATGTCGCTGTGATTT	120
Db	580	AGGCAGAGTTTGGACGAGAGAAATTTGAAATTTGCTGAACAGAANAATGCCGCAATTGATGG	639
QY	121	CACCTCAGGAAACGTGCTCAGGGGGGAGAACCCCTTGCGTGTCTAAATAGTGGCGCTGA	180
Db	640	CTTTGAGGAAGAGAGCTCAAGAGAGAAAGCCTTTGGCTGAGGCCAAATCGTGGGTTGCA	699
QY	181	CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGGGGGCTCAGT	240
Db	700	CACACATCACTGCTCAGACTGCTGTGCTTATGAAACTCTGGGTGCTCGGGGGGCCAGT	759
QY	241	GCCGCTGGTCTGCTTGTAAACATCTACTCAACTCAGATGAAGTAGCTGCAGCACTGGCTG	300
Db	760	GCCGATGGGCTGCCCTGCACACATCTATCCACTCTCATGAAGTGGCTGCTCTAGCAG	819
QY	301	AGGCTGGAGTTGCAGTGTTCGCTTGGAAAGGGCGAGTCAGAAGATGACTTCTGGTGTGTA	360
Db	820	AAAGTGGATTTCCCTGTTTTGCTTGGAAAGGAGAGATCAGAAGATGACTTTTGGTGTGTA	879
QY	361	TTGACCCGCTGTGTGAACATGCATGGGCTGGCAGGCCAACATGATCCTGATGATGGGG	417
Db	880	TCGATAGATGTGTGAATGTGAGGGCTGGCAGCCCAACATGATCTTGGATGATGGAG	936

RESULT 11	
AB020635	
LOCUS	AB020635
DEFINITION	5025 bp mRNA linear PRI 16-JUN-1999
ACCESSION	Homo sapiens mRNA for KIAA0828 protein, partial cds.
VERSION	AB020635
KEYWORDS	AB020635.1 GI:4240144
SOURCE	Homo sapiens adult male brain cDNA to mRNA, clone_11b:pbuescript11
	SK plus clone:nh04230.

ORGANISM Homo sapiens;
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 ¹⁶/₁₆ (sites)

MIYAJIMA, N., TANAKA, A., KOTANI, H., NOMURA, N. and OHARA, O.
Prediction of the coding sequences of unidentified human c

XII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro
DNA Res. 5 (6), 355-364 (1998)

JOURNAL MEDLINE 99156230
REFERENCE 2 (bases 1 to 5025)
AUTHORS Ohara, O., Suyama, M., Kikuno, R., Nagase, T. and Ishikawa, K.
TITLE Direct Submission
JOURNAL Submitted (02-DEC-1998) Osamu Ohara, Kazusa DNA Research Institute,

FEATURES
source Laboratory of DNA Technology; Yana 1532-3, Kisarazu, Chiba 292-0812, Japan (E-mail: cdnainfo@kazusa.or.jp, Tel: +81-438-52-3913, Fax: +81-438-52-3914)
Location/Qualifiers I. 5025

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="hh04230"
/sex="male"
/tissue_type="brain"
/clone_1lb="pbluescriptII SK plus"
/dev_stage="adult"
1. .5025
/gene="KIAA0828"
<1. .1860
CDS

```

```

CDS
        /gene="KIAA0828"
        <1..1860
        /gene="KIAA0828"
        /codon_start=1
        -/product="KIAA0828 protein"
        /protein_id="BAA74851.1"
        /db_xref="GI:4240145"
        /translation="EPVVAAEAVMVSQVVSAAAAAKVPEVELKDLSPSEAESQIGLST
AAVGAMAPRAGGDPDEAPAPAPAEERPVPVPGSPGSPAALSPAAGKVPQASAKRKSDPHH
OHQRHRDGGEGALVSPDGTVTETAPRTVKIKIDFADOKQEFNKRPTKIGRISRSISQSS
TDSYSSAASYTDSDDDETSPRDKQKQNSKSSDFCVKNIKQAEFGREIETAEQEMPA
LMALRKRAOGEKPLAGAKIVGCTHITAQFANVLEMTLGAAGOCRWACNNTSTINEVA
AALAESGFVPFAWKGESEDDFWWCIDRCVAVNEGQPMNMLIDDGDLTHWITKKYPNMF
KKIKGIVEESVTGVHRLYLQLSKAGKLCVPMANVNDSVTKQKEDNLCCRSILDGKLR
TTDMFEGGKQVYVCGYGEVGKGCALKAMGSIYVTEIDPICALQACDGFRLVKLN
EVIKQVDIVITCTGNKNVTVREHLDRMKNSCIVCNMGHSNTEIDVASLKPETLWERV
RSQVDHVIWPDGKRIVLLAEGRLNLSCSYVTEFLSITATQALALIEYNAPEGRY
KQDVYLLPKKNDIYVASLHLPTFDAHLTELTEQAKYLGLNKNGPFKPNYYRY"
BASE COUNT      1281 a      1179 c      1136 g      1429 t
ORIGIN

```

Query Match.	65.9%;	Score 274.6;	DB 9;	Length 5025;
Best Local Similarity	78.7%;	Pred. No. 2.5e-73;		
Matches 328;	Conservative	0;	Mismatches 89;	Indels 0;
				Gaps 0

QY	1	GAGAGACGACGCAAAACCCTCCAAAGGGCAGCAGCAATTTCTGTGTGAAGAACAATCAAGC	60
Db	542	GGGACAGCAGCAAAAAGAACTTAAGGGAGCAGTACTTCTGTGTTAAGAACATCAAAAC	6011
QY	61	AGGCAGAAATTTGAGCGCCGGAGATTGAGATTTGCAGAGCAAGACATGTCTGCTGATTT	120
Db	602	AGGCAGAGTTTGGACGAGAGAAATTGAAYTTGCTGAACAAGAAATGCCCTGCATTTGATGG	6611
QY	121	CACTCAGAAACGTGCTCAGGGGAGAAAGCCCTTGCTGTGCTAAATAGTGGCTGTA	180
Db	662	CTTTGAGGAAGAGAGAGCTCAAGAGAGAAAGCCCTTTGGCTGAGCCAAAATCGTGGGTTGCA	7211
QY	181	CACACATCACAGCCCCAGACAGCGGTGTGATTTGAGACACTCTGTGCCCTGGGGGCTCAGT	240
Db	722	CACACATCACATGCTCTCAGACTGCTGTGCTTATGGAACAATCTGGGTGCTCTGGGGGCCAGT	7811
QY	241	GCCGCTGCTGTGCTTGTACATCTACTCACTCAGATGAAGTAGCTGCAGCATGGCTG	300
Db	782	GCCGATGGGCTGCGCTGCACACATCTATTCACCTCTCAATGAAGTGGCTGCTGCTCTAGCAG	8411
QY	301	AGGCTGGAGTTGCAGTGTTCGCTTTGGAAGGGCGAGTCAGAAAGATGACTTCTGGTGGTGA	360
Db	842	AAAGTGGATTTCTCTGTTTTTGGCCTGGAGGGAGAGTCAAGAGATGACTTTTGGTGGTGA	9011
QY	361	TTGACCGCTGTGTGAACATGGATGGGTGGCAGGGCCAACATGATCCTGGATGATGGGG	417
Db	902	TGCAATAGATGTGTGAATGTGGAGGGCTGGCAGCCCAACATGATCTTGGATGATGGAG	958

RESULT 12					
BC024325					
LOCUS	BC024325	5052 bp	mrna	linear	PRI 12-MAR-2002
DEFINITION	Homo sapiens, KIA00828 protein, clone MGC:21525 IMAGE:3907552,				
	mrna, complete cds.				
ACCESSION	BC024325				

VERSION	BC024325.1	GI:19353091
KEYWORDS	MG.	
SOURCE	Homo sapiens.	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo;	
REFERENCE	1 (bases 1 to 5052)	

AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK	NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT	Contact: MGC help desk

Email: cgaphs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcdéparali.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://limage.llnl.gov>
Series: IRAC Plate: 22 Row: 1 Column: 22
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

```

FEATURES      Location/Qualifiers
source        1. .5052

```

```

/organism="Homo sapiens"
/db_xref="locusID:23382"
/db_xref="taxon:9606"
/clone="MGC:21525 IMAGE:3907552"
/tissue_type="Uterus, leiomyosarcoma"
/clone_1lb="NIH_MGC_71"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
55. .1890

```

```

/codon_start=1
/product="K1AA0828 protein"
/protein_id="AAH24325.1"
/db_xref="GI:19353092"
/translation="MSVQVVSAAAAAKVPEVELKDLSPSEAESQLGLSTAAYGAMAPP
AGGDPPEAPAPAERPPVPGPGSGPALSLSPAAGKVPQASAMKRSDPHHQHRHRDGG
EALVSPDGTVEAPRTVKQIQFADQKQEFKRPRTIGRSLRSISQSSSTDSYSSAA
SYTDSDDETSPRDKQKQKSSGSDCYVKNIKQAEFGREIEIAEOEMPALMLRKRA
OGKPLAGAKIVGCTHITAOATSLMETGLGAOCRWAAACNISTLNEVAALAESEGF
PVFAWKGESEDDFWWCIDRCVNVGEWQPMNLLDDGGDLTHWIIKKYPNMFKIKGIVEE
ESVTGVHRLYQLSKAGKLCVPAMVNDSTYRKQEDNLCCRESILDLGKRTDMMFEGG
KQVVVCGYGEVGGCCALKAMGSIYVYTEIDPICALQACMDGFRLYKLNIEIRQVDII
VITCTGNKNVYTRHLDKMSNCSTVPTFVLSITATQALAILELYNAPEGRYKQDYYLLP
WPDGKRIYLLAEGRLNLDKMSNCSTVPTFVLSITATQALAILELYNAPEGRYKQDYYLLP
KKMDEYVASLHLPTFDAHLTELTDQAKYLGUNKNGPEKPNYYRY"

```

BASE COUNT	1300 a	1184 c	1145 g	1423 t
ORIGIN				

Query Match	65.98;	Score 274.6;	DB 9;	Length 5052;
Best Local Similarity	78.7%;	Pred. No. 2.5e-73;		
Matches 328; Conservative	0;	Mismatches 89;	Indels 0;	Gaps 0;

[illegible]

QY	121	CACATCAGGA	AACGCTGCTCAGGGGGAGAACCCCTTGCGCTGTCTAAATAGTGGCGTGA	180
Db	692	CTTTGAGGA	GAGAGAGCTCAAGGAGAAAAAGCCTTTGGCTGGAGCCAAAATCGTGGGTTGCA	751
QY	181	CACACATCACAGCCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGGGGGCTCAGT	240	
Db	752	CACACATCACATGCTCTCAGACTGTGTCTATTGGAACCTCGGGTGTCTGGGGGCCACGT	811	
QY	241	GCCGCTGCTCTGCTTGTACATCTACTCAACTCAGAATGAAGTAGCTGCAGACACTGSGCTG	300	
Db	812	GCCGATGGGCGCTGCCTGCACACATCTATTCCACTCTCATAGAAGTGGCTGCTGCTCFACAG	871	
QY	301	AGGCTGAGCTTGCAGTGTTCGCTTGGAAAGGGCGAGTCAGAAGATGACTTCTGCTGTGTA	360	
Db	872	AAAGTGAATTTCCCTGTGTTTTTGGCCTGGAAAGGAGAGTCAGAAGATGACTTTTGGTGGTGTGA	931	
QY	361	TTCAGCGCTGTGTGAACATGGATGGGTGGCAGGCCAACATGATCCTGGATGATGGGG	417	
Db	932	TTCGATGATGTGTGAATGTGTGAGGGGCTGGCAGCCCAAACATGATCTTGGATGATGAG	988	

RESULT 13	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
AY113501	AY113501	Drosophila melanogaster	1792 bp mRNA	linear	full insert	CDNA.	INV 21-MAY-2002
AY113501	AY113501	Drosophila melanogaster	1792 bp mRNA	linear	full insert	CDNA.	INV 21-MAY-2002
AY113501.1	AY113501.1	Drosophila melanogaster	1792 bp mRNA	linear	full insert	CDNA.	INV 21-MAY-2002
FLI_CDNA.	FLI_CDNA.	Drosophila melanogaster	1792 bp mRNA	linear	full insert	CDNA.	INV 21-MAY-2002
Drosophila melanogaster	Drosophila melanogaster	Drosophila melanogaster	1792 bp mRNA	linear	full insert	CDNA.	INV 21-MAY-2002

REFERENCE
AUTHORS
1 (bases 1 to 1792)
Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Carlson, J.,

Champe, M., Chavez, C., Dorsett, V., Dresnek, D., Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Krommiller, B., Li, P., Liao, G., Miranda, A., Mungall, C. J., Nuno, J., Pacle, J., Paragas, V., Park, S., Patel, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Rubin, G. M. and Celniker, S.

TITLE Direct Submission
JOURNAL Submitted (16-MAY-2002) Berkeley Drosophila Genome Project,
Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, USA
COMMENT Sequence submitted by:

Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720

This clone was sequenced as part of a high-throughput process to sequence clones from Drosophila Gene Collection 1 (Rubin et al., Science 2000). The sequence has been subjected to integrity checks for sequence accuracy, presence of a polyA tail and contiguity within 100 kb in the genome. Thus we believe the sequence to reflect accurately this particular cDNA clone. However, there are artifacts associated with the generation of cDNA clones that may have not been detected in our initial analyses such as internal priming, priming from contaminating genomic DNA, retained introns due to reverse transcription of unspliced precursor RNAs, and reverse transcriptase errors that result in single base changes. For further information about this sequence, including its location and relationship to other sequences, please visit our web site (<http://fruitfly.berkeley.edu>) or send email to cdna@fruitfly.berkeley.edu.

```

FEATURES
source
Location/Qualifiers
1..1792
/organism="Drosophila melanogaster"
/strain="Y; cn bw sp"
/db_xref="taxon:7227"
/map="89E10-89E10"
1..1792
/gene="Ancy89E"
/note="alignment with genomic scaffold AE003715"
/db_xref="FlyBASE:FBgn0015011"

```


TITLE The morphogenetic and regulatory functions of the Drosophila Abdominal-B gene are encoded in overlapping RNAs transcribed from separate promoters

JOURNAL Genes Dev. 3 (12A), 1969-1981 (1989)

REFERENCE PUBMED 2482824
AUTHORS 4 (bases 1 to 80423)
Martin, C.H., Celniker, S.E., Davis, C.A., Mayeda, C.A., Strathmann, M.P., Yoshida, K. and Palazzolo, M.J.

JOURNAL Unpublished (1992)
FEATURES Location/Qualifiers
source 1..80423
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/map="3, 89E; polytene"
/tissue_type="whole fly"
/dev_stage="adult"
/tissue_1lb="p1"

misc_feature 1..80423
BASE COUNT 23439 a 16747 c 16641 g 23596 t
ORIGIN

Query Match 45.8%; Score 191; DB 3; Length 80423;
Best Local Similarity 66.3%; Pred. No. 1.5e-47;
Matches 275; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 3 GAGAAGCAGCAAAACCACTCCAGAGGCGAGACCAATTCTGTGTGAAGACATCAAGCAG 62
DB 3342 GACAACCATCAGCGCACTCCGCCGAGGACACTTTCGCGTGAAGACATCTCGAAG 3401
QY 63 GCAGAAATTGGACGCCGGAGATGAGATTGCAGAGCAAGACATGCTGCTGATTCA 122
DB 3402 AGTGCAATTGGAAAGCGGAGATCGAGATCGCCGAGTCGAGATGCCGGCATCATGACT 3461
QY 123 CTCAGGAAACGTGCTCAGGGGAGAGCCCTTGCTGCTAAATAGTGGGCTGTACA 182
DB 3462 CTGAGGAAAGAGCGAAGATGAGAGCCCTAAAGGGTGCCAATATGTCGATGCACC 3521
QY 183 CACATCACAGCCCGAGACAGCGGTGTTGATTGAGACACTCTGTGCCCTGGGGCTCAGTGC 242
DB 3522 CACGTCAATGCTCAGTCGCGAGTGTGATCGAGACCCCTGCTCAACTGGGCGCCACAGTT 3581
QY 243 CGCTGCTCTGCTTGTACATCTACTCAACTCAGAATGAGTAGCTGACACTGGCTGAG 302
DB 3582 CGCTGGGCTGCTGCAACATTTATCCACAAACGAGTGCCTGCTGCGCTGCGCAGAG 3641
QY 303 GCTGAGTTGCACTGTTCGCTTGAAGCGGAGTCAAGAGATGACTTCTGCTGTATTT 362
DB 3642 GCGGGAATTCGATCTTCGCTTGGCGGAGAGAGAGAGAGAGAGTCTGCTGCTTG 3701
QY 363 GACCGCTGTGTGAACATGATGGGTGGCAGGCCAATGATCTGTGATGATGGG 417
DB 3702 GACAGGGCCATCTACTCCGACGGCTGGCAGCCGAACCTAATCTGAGACGCGCG 3756

Search completed: April 21, 2003, 02:39:04
Job time : 1107.44 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2003, 23:48:04 ; Search time 18.6401 Seconds
(without alignments)
6860.687 Million cell updates/sec

Title: US-09-782-051-1_COPY_529_945

Perfect score: 417
Sequence: 1 gagagaagcagcaaaccaac.....catgatcctgatgatggg 417

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/lna/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/lna/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/lna/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/lna/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/lna/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/1/lna/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
c 1	226.4	54.3	636	4	US-09-328-111-850
2	138.8	33.3	2211	4	US-09-318-448-26
3	138.8	33.3	2211	4	US-09-347-878-2
4	118.2	28.3	1812	2	US-08-669-536-1
5	101.8	24.4	1767	3	US-08-930-894-1
6	72.6	17.4	289	1	US-08-204-740-8
7	72.6	17.4	289	3	US-09-081-167A-8
8	72.6	17.4	289	3	US-09-081-395-8
9	72.6	17.4	289	4	US-09-416-833-8
10	72.6	17.4	289	5	PCT-US95-02521-8
11	68.4	16.4	285	1	US-08-204-740-6
12	68.4	16.4	285	3	US-09-081-167A-6
13	68.4	16.4	285	3	US-09-081-395-6
14	68.4	16.4	285	4	US-09-416-833-6
15	68.4	16.4	285	5	PCT-US95-02521-6
c 16	33.8	8.1	7218	1	US-08-232-463-14
17	32.8	7.9	31571	1	US-08-323-443B-1
18	32.8	7.9	53526	3	US-08-658-136-2
19	32.8	7.9	53577	3	US-08-658-136-1
20	31.2	7.5	233	2	US-08-687-080-108
21	31	7.4	3306	4	US-09-770-170-7
22	30.2	7.2	1317	4	US-09-134-078-2
c 23	29.4	7.1	4483	4	US-08-961-527-363
24	29.4	7.1	12394	4	US-09-488-856A-10
c 25	29	7.0	1803	1	US-08-021-608D-7
c 26	29	7.0	1803	1	US-08-726-160-7
c 27	29	7.0	1803	5	PCT-US94-01782-7

c 28	29	7.0	2381	1	US-08-021-608D-9	Sequence 9, Appl1
c 29	29	7.0	2381	1	US-08-726-160-9	Sequence 9, Appl1
c 30	29	7.0	2381	5	PCT-US94-01782-9	Sequence 9, Appl1
c 31	29	7.0	2384	1	US-08-021-608D-1	Sequence 1, Appl1
c 32	29	7.0	2384	1	US-08-726-160-1	Sequence 1, Appl1
c 33	29	7.0	2384	5	PCT-US94-01782-1	Sequence 1, Appl1
c 34	28.8	6.9	18627	4	US-08-961-527-113	Sequence 113, App
c 35	28.6	6.9	3744	4	US-08-961-527-263	Sequence 263, App
c 36	28.6	6.9	5816	4	US-09-220-641-4	Sequence 4, Appl1
c 37	28.2	6.8	3879	4	US-08-916-352-1	Sequence 1, Appl1
c 38	28.2	6.8	3979	4	US-08-085-199B-10	Sequence 10, Appl
c 39	28.2	6.8	6827	4	US-08-961-527-60	Sequence 60, Appl
c 40	28	6.7	308	1	US-08-627-706-5	Sequence 5, Appl1
c 41	28	6.7	308	4	US-09-103-489-5	Sequence 5, Appl1
c 42	28	6.7	500	1	US-08-627-706-9	Sequence 9, Appl1
c 43	28	6.7	500	4	US-09-103-489-9	Sequence 9, Appl1
c 44	28	6.7	1875	4	US-09-761-716-1	Sequence 1, Appl1
c 45	27.8	6.7	1611	2	US-08-650-598-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-09-328-111-850/c

; Sequence 850, Application US/09328111
; Patent No. 6262333

GENERAL INFORMATION:

APPLICANT: Endege, Wilson O.
APPLICANT: Steilmann, Kathleen E.
APPLICANT: Astle, Jon H.
APPLICANT: Burgess, Christopher C.
APPLICANT: Bushnell, Steven E.
APPLICANT: Carroll III, Eddie
APPLICANT: Catlino, Theodore J.
APPLICANT: Dertl, Adnan
APPLICANT: Ford, Donna M.
APPLICANT: Lewis, Marcia E.
APPLICANT: Monahan, John E.
APPLICANT: Schlegel, Robert
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
TITLE OF INVENTION: PRODUCTS
FILE REFERENCE: CCD-257 (US)
CURRENT APPLICATION NUMBER: US/09/328,111
EARLIER FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: US 60/088,801
NUMBER OF SEQ ID NOS: 850
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 850
LENGTH: 636
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(636)
OTHER INFORMATION: n = A,T,C or G
US-09-328-111-850

Query Match 54.3%; Score 226.4; DB 4; Length 636;
Best Local Similarity 99.2%; Pred. No. 4e-65;
Matches 238; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY	179	TACACACATCAGAGCCAGACAGC-GGTGTGATGTAGACACTCTGCCCCGGGGCTC	237
DB	420	TACACACATCAGAGCCAGACAGC-GGTGTGATGTAGACACTCTGCCCCGGGGCTC	361
QY	238	AGTGCCGCTGCTGCTGTGATCACTCACTCAGATGAAGTGAAGCTGCAGCACTGG	297
DB	360	AGTGCCGCTGCTGCTGTGATCACTCACTCAGATGAAGTGAAGCTGCAGCACTGG	301
QY	298	CTGAGGCTGAGTGTGCTGCTGCTGGAAGGCGAGTCAGAGTGAAGTGAAGTGTGCTG	357

Db 300 CTGAGGCTGAGTTGCAAGTTCGCTTGGAGGCGGAGTCAGAGATGACTTCTGCTGT 241
QY 358 GTATTGACCGCTGTGTGACATGATGGGTGGAGGCCCAACATGATCTGTGATGGGG 417
Db 240 GTATTGACCGCTGTGTGACATGATGGGTGGAGGCCCAACATGATCTGTGATGGGG 181

RESULT 2

US-09-318-448-26
; Sequence 26, Application US/09318448
; Patent No. 6210950
; GENERAL INFORMATION:
; APPLICANT: Johnson, William G.
; APPLICANT: Stenroos, Edward S.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; TITLE OF INVENTION: DEVELOPMENTAL DISORDERS
; FILE REFERENCE: 601-1-057
; CURRENT APPLICATION NUMBER: US/09/318,448
; CURRENT FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; LENGTH: 2211
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-318-448-26

Query Match 33.3%; Score 138.8; DB 4; Length 2211;
Best Local Similarity 61.2%; Pred. No. 5.7e-36;
Matches 224; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 52 ACATCAAGCAGCAGAAATTTGGAGCGCGGAGATTGAGATTGACAGCAAGCATGTCTG 111
Db 79 ACATCGCGCTGGCTGGCTGGGAGCAGAGCGCTGGACATTGTGAGACGAGATGCCGG 138
QY 112 CTCTGATTCTACTGAGAAACGTGCTCAGGGGAGAGACCCCTTGCTGTGCTAAATAG 171
Db 139 GCCTGATGCGTATGCGGAGAGCGGTACTCGGCTCCAAAGCCACTGAAGGGCGCCGATCG 198
QY 172 TGGCTGTACACATCAACAGCCAGACAGCGGTGTGATTGAGACACTGTGCCCTGG 231
Db 199 CTGGCTGCCGTGACATGACCGTGGAGACGCGGCTCTCATTTGAGACCCCTGCACCCCTGG 258
QY 232 GGGCTCAGTCCGCTGCTGCTGTGTAACATCTACTCACTCAGATGAGTACTGACG 291
Db 259 GTGCTGAGGTGAGTGTCCAGCTGCAACATCTTCTCCACCAGAACCATGGCGGCTG 318
QY 292 CACTGGCTGAGCTGAGTGTGAGTGTTCGCTTGGAGGGCGAGTCAAGATGACTTCT 351
Db 319 CCATTGCCAAGGCTGAGTTCGCTGTATGCTTGAAGGGCGAAGGAGAGAGTACC 378
QY 352 GGTGTGTATTGACCGCTGTGTGAACATGATGGGTGGAGGCCCAACATGATCTGTGATG 411
Db 379 TGTGTGATGAGCAGACACCTGTACTTCAAGAGAGCGGCCCTCAACATGATCTGTGAGC 438
QY 412 ATGGGG 417
Db 439 ACGGGG 444

RESULT 3
US-09-347-878-2
; Sequence 2, Application US/09347878C
; Patent No. 6376210
; GENERAL INFORMATION:
; APPLICANT: Yuan, Chong
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES
; FILE REFERENCE: 25885-1651
; CURRENT APPLICATION NUMBER: US/09/347,878C
; CURRENT FILING DATE: 1999-07-06
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2

; LENGTH: 2211
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Human S-adenosylhomocysteine hydrolase cDNA
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: M61831/GenBank
US-09-347-878-2

Query Match 33.3%; Score 138.8; DB 4; Length 2211;
Best Local Similarity 61.2%; Pred. No. 5.7e-36;
Matches 224; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 52 ACATCAAGCAGCAGAAATTTGGAGCGCGGAGATTGAGATTGACAGCAAGCATGTCTG 111
Db 79 ACATCGCGCTGGCTGGCTGGGAGCAGAGCGCTGGACATTGTGAGACGAGATGCCGG 138
QY 112 CTCTGATTCTACTGAGAAACGTGCTCAGGGGAGAGACCCCTTGCTGTGCTAAATAG 171
Db 139 GCCTGATGCGTATGCGGAGCGGTACTCGGCTCCAGCCACTGAAGGGCGCCGATCG 198
QY 172 TGGCTGTACACATCAACAGCCAGACAGCGGTGTGATTGAGACACTGTGCCCTGG 231
Db 199 CTGGCTGCCGTGACATGACCGTGGAGAGCGGCGCTCTCATTTGAGACCCCTGTCACCCCTGG 258
QY 232 GGGCTCAGTCCGCTGCTGCTGTGTAACATCTACTCACTCAGATGAGTACTGACG 291
Db 259 GTGCTGAGGTGAGTGTCCAGCTGCAACATCTTCTCCACCAGAACCATGGCGGCTG 318
QY 292 CACTGGCTGAGCTGAGTGTGAGTGTTCGCTTGGAGGGCGAGTCAAGATGACTTCT 351
Db 319 CCATTGCCAAGGCTGAGTTCGCTGTATGCTTGAAGGGCGAAGGAGAGAGTACC 378
QY 352 GGTGTGTATTGACCGCTGTGTGAACATGATGGGTGGAGGCCCAACATGATCTGTGATG 411
Db 379 TGTGTGATGAGCAGACACCTGTACTTCAAGAGAGCGGCCCTCAACATGATCTGTGAGC 438
QY 412 ATGGGG 417
Db 439 ACGGGG 444

RESULT 4
US-08-669-536-1
; Sequence 1, Application US/08669536
; Patent No. 5910444
; GENERAL INFORMATION:
; APPLICANT: MASUTA, CHIKARA
; APPLICANT: UEHARA, KYOKO
; APPLICANT: TANAKA, HIDEO
; APPLICANT: KAWATA, SHIGERU
; TITLE OF INVENTION: ORGANISMS IN WHICH THE EXPRESSION OF
; TITLE OF INVENTION: S-ADENOSYLHOMOCYSTEINE HYDROLASE GENE IS INHIBITED
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/669,536
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: MURPHY JR, GERALD M
; REGISTRATION NUMBER: 28,977


```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/204,740
; FILING DATE: 04-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: NO. 5753432nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 93,354-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 289 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
;
US-08-204-740-8

```

Query Match	17.4%	Score 72.6;	DB 1;	Length 289;
Best Local Similarity	58.6%	Pred. No. 1.4e-14;		
Matches 126;	Conservative 0;	Mismatches 89;	Indels 0;	Gaps 0;
QY 52	ACATCAAGCAGGCGAGAAATTTGCACGCCGGAGATTGAGATTGCAGACCAACATGTCG	111		
DB 75	ACATCGGCCTGGCTGCCTGGGGACGCCAAGGCCCTGCACATTTGCTGAGAACGAGATGCCGG	134		
QY 112	CTCTGATTTCACTCAGGAACGTCCTCAGGGGGAGAACCCCTTGCTGCTAAATAG	171		
DB 135	GCCTGATGCGTATGCGGGAGCGGTACTCGGCCCTCCAAGCCACTGAGAGGGCCCGCATCG	194		
QY 172	TGGGCTGTACACACATCACAGCCACAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGG	231		
DB 195	CTGGCTGCCTGCACACATGACCCGTGGAGACGGCCGTCCTCATTTGAACCCCTGTCACCCCTGG	254		
QY 232	GGGCTCAGTGCCCGCTGGTCTGCTTGTACATCTAC	266		
DB 255	GTCTGTAGAGGTGACAGTGTCTCCAGCTGCAACATCTTC	289		

RESULT 7
US-09-081-167A-8
; Sequence 8, Application US/09081167A
; Patent No. 6083745
; GENERAL INFORMATION:
; APPLICANT: Gudkov, Andrei
; APPLICANT: Kazarov, Alexander
; APPLICANT: Mazo, Ilya
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods for Identifying Genetic
; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
; TITLE OF INVENTION: Growth In Cancer Cells
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 S. Wacker Drive, 32nd Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/081,167A

```

? FILING DATE: 18-MAY-1998
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: NO. 6083745nan, Kevin E
? REGISTRATION NUMBER: 35,303
? REFERENCE/DOCKET NUMBER: 93,354-KK
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 312-913-0001
? TELEFAX: 312-913-0002
? TELEX:
? INFORMATION FOR SEQ ID NO: 8:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 289 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: CDNA

```

Query Match	17.48;	Score 72.6;	DB 3;	Length 289;
Best Local Similarity	58.68;	Pred. No. 1.4e-14;		
Matches 126;	Conservative 0;	Mismatches 89;	Indels 0;	Gaps 0;
QY 52	ACATCAAGCAGCGCAGAAATTGGACGCCGGAGATTGAGATTGCAGACCAAGCATGTCTG	111		
Db 75	ACATCGGCCTGGCTGCGCTGGGGACGCCAAGGCCCTGACATTGCTGAGAACGAGATGCCGG	134		
QY 112	CTCTGATTTCACTCAGGAACGTGCTCAGGGGGAGAAGCCCTTGCTGTGCTAAATAG	171		
Db 135	GCGTGATGCGTATATGCGGAGCGGTACTCGCCCTCCAAGCCACTGAAGGGCCCGCATCG	194		
QY 172	TGGGCTGTACACACATCACAGCCCAAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGG	231		
Db 195	CTGGCTGCCTGCACATGACCCGTGGAGACGGCCGCTCCTCATTTGAGACCCCTGTCACCCCTGG	254		
QY 232	GGGCTCAGTGCCGCTGGTCTGCTTGTAAACATCTAC	266		
Db 255	GTGCTGAGGTGACAGTGTGCTCCAGCTGCACACATCTTC	289		

RESULT 8
 US-09-081-395-8
 Sequence 8, Application US/09081395
 Patent No. 6083746
 GENERAL INFORMATION:
 APPLICANT: Gudkov, Andrei
 APPLICANT: Kazarov, Alexander
 APPLICANT: Mazo, Ilya
 APPLICANT: Roninson, Igor B
 TITLE OF INVENTION: Methods for Identifying Genetic
 TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
 TITLE OF INVENTION: Growth in Cancer Cells
 NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
 STREET: 300 S. Wacker Drive, 32nd Floor
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/081,395
 FILING DATE: 18-MAY-1998
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: NO. 6083746nan, Kevin E
 REGISTRATION NUMBER: 35,303
 REFERENCE/DOCKET NUMBER: 93,354-KK

QY 232 GGGCTCAGTCCCGCTGCTGCTTGTACATCTAC 266
| | | | | | | | | | | | | | | | | | | | | |
DB 255 GTGCTGAGTGCAGTGTCCAGCTGCACATCTTC 289

RESULT 11

US-08-204-740-6

; Sequence 6, Application US/08204740

; Patent No. 5753432

; GENERAL INFORMATION:

; APPLICANT: Gudkov, Andrei

; APPLICANT: Kazarov, Alexander

; APPLICANT: Mazo, Ilya

; APPLICANT: Roninson, Igor B

; TITLE OF INVENTION: Methods for Identifying Genetic

; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant

; TITLE OF INVENTION: Growth in Cancer Cells

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Allgretti & Witcoff, Ltd.

; STREET: 10 S. Wacker Drive, Suite 3000

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/204,740

; FILING DATE: 04-MAR-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: NO. 5753432nan, Kevin E

; REGISTRATION NUMBER: 35,303

; REFERENCE/DOCKET NUMBER: 93,354-C

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 312-715-1000

; TELEFAX: 312-715-1234

; TELEX: 910-221-5317

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 285 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: CDNA

; US-08-204-740-6

Query Match 16.4%; Score 68.4; DB 1; Length 285;
Best Local Similarity 56.8%; Pred. No. 3.4e-13;
Matches 126; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 45 GTGAAGACATCAAGCAGGAGAAATTGGACGCCGGAGATTGAGATTGCAGAGCAGAC 104
| | | | | | | | | | | | | | | | | | | | | |
DB 64 GTCCGGACATCGAGTGGCCGCTGGGAGGAGAGCTCTGATATAGCTGAGATGAG 123
| | | | | | | | | | | | | | | | | | | | | |
QY 105 ATGCTGCTCTGATTCTACTCAGAAACGTGCTCAGGGGAGAGAGCCCTGGCTGCT 164
| | | | | | | | | | | | | | | | | | | | | |
DB 124 ATGCGAGGTTGATGCGCATGCGGAGATGATCTCAGCCTCCAGCCACTGAAGGCTGCT 183
| | | | | | | | | | | | | | | | | | | | | |
QY 165 AAATAGTGGCTGTACACATCAGAGCCAGAGAGCGGTGTGATTGAGACACTCTGT 224
| | | | | | | | | | | | | | | | | | | | | |
DB 184 CGCATTTGCTGCTGCTGCGCATGACCGTGAGAGACTGCTCTCATTGAGACTCTGCTG 243
| | | | | | | | | | | | | | | | | | | | | |
QY 225 GCCCTGGGGCTCAGTGGCGCTGCTGCTTGTACATCTAC 266
| | | | | | | | | | | | | | | | | | | | | |
DB 244 GCCCTGGGTGCTGAGGCGCGGTGCTCAGCTGCAACATCTTC 285
| | | | | | | | | | | | | | | | | | | | | |

RESULT 12

US-09-081-167A-6
; Sequence 6, Application US/09081167A

; Patent No. 6083745

; GENERAL INFORMATION:

; APPLICANT: Gudkov, Andrei

; APPLICANT: Kazarov, Alexander

; APPLICANT: Mazo, Ilya

; APPLICANT: Roninson, Igor B

; TITLE OF INVENTION: Methods for Identifying Genetic

; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant

; TITLE OF INVENTION: Growth in Cancer Cells

; NUMBER OF SEQUENCES: 15

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff

; STREET: 300 S. Wacker Drive, 32nd Floor

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/081,167A

; FILING DATE: 18-MAY-1998

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: NO. 6083745nan, Kevin E

; REGISTRATION NUMBER: 35,303

; REFERENCE/DOCKET NUMBER: 93,354-KK

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 312-913-0001

; TELEFAX: 312-913-0002

; TELEX:

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 285 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: CDNA

; US-09-081-167A-6

Query Match 16.4%; Score 68.4; DB 3; Length 285;
Best Local Similarity 56.8%; Pred. No. 3.4e-13;
Matches 126; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 45 GTGAAGACATCAAGCAGGAGAAATTGGACGCCGGAGATTGAGATTGCAGAGCAGAC 104
| | | | | | | | | | | | | | | | | | | | | |
DB 64 GTCCGGACATCGAGTGGCCGCTGGGAGGAGAGCTCTGATATAGCTGAGATGAG 123
| | | | | | | | | | | | | | | | | | | | | |
QY 105 ATGCTGCTCTGATTCTACTCAGAAACGTGCTCAGGGGAGAGAGCCCTGGCTGCT 164
| | | | | | | | | | | | | | | | | | | | | |
DB 124 ATGCCAGGTTGATGCGCATGCGGAGATGATCTCAGCCTCCAGCCACTGAAGGCTGCT 183
| | | | | | | | | | | | | | | | | | | | | |
QY 165 AAATAGTGGCTGTACACATCAGAGCCAGAGAGCGGTGTGATTGAGACACTCTGT 224
| | | | | | | | | | | | | | | | | | | | | |
DB 184 CGCATTTGCTGCTGCTGCGCATGACCGTGAGAGACTGCTCTCATTTGAGACTCTGCTG 243
| | | | | | | | | | | | | | | | | | | | | |
QY 225 GCCCTGGGGCTCAGTGGCGCTGCTGCTTGTACATCTAC 266
| | | | | | | | | | | | | | | | | | | | | |
DB 244 GCCCTGGGTGCTGAGGCGCGGTGCTCAGCTGCAACATCTTC 285
| | | | | | | | | | | | | | | | | | | | | |

RESULT 13

US-09-081-395-6

; Sequence 6, Application US/09081395

; Patent No. 6083746

; GENERAL INFORMATION:

; APPLICANT: Gudkov, Andrei

; APPLICANT: Kazarov, Alexander

APPLICANT: MAZO, Ilya
 TITLE OF INVENTION: Roninson, Igor B
 TITLE OF INVENTION: Methods for Identifying Genetic
 TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
 TITLE OF INVENTION: Growth in Cancer Cells
 NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
 STREET: 300 S. Wacker Drive, 32nd Floor
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/081,395
 FILING DATE: 18-MAY-1998
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: NO. 6083746nan, Kevin E
 REGISTRATION NUMBER: 35,303
 REFERENCE/DOCKET NUMBER: 93,354-KK
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312-913-0001
 TELEFAX: 312-913-0002
 TELEX:
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 285 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: CDNA

```

Query Match      16.4%; Score 68.4; DB 3; Length 285;
Best Local Similarity 56.8%; Pred. No. 3.4e-13;
Matches 126; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 45 GTGAGACATCAAGCAGCAGCAATTTGGACGCCGGAGATTGAGATTGCAGACGACAGAC 104
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 64 GTGCGGACATCGGACTGGCCGCTGGGGACGGAAGCTCTGATATAGCTGAGATGAG 123
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 105 ATGCTGCTCTGATTTCTACTCAAGAAACGTGCTCAGGGGAGAAAGCCCTGGCTGCT 164
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 124 ATGCCAGGCTTGATGCGCATGCGGAGAGATGTACTACGCTCCAAAGCCACTGAGGGTGCT 183
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 165 AAATAGTGGGCTGTACACACATCAAGCCAGACAGCGGTGTGATTGAGACACTCTGT 224
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 184 CGCATGTGCTGCTGCGCTGCGCATGACCGGTGAGAGACTGTGTCTATTGAGACTCTCGTG 243
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 225 GCCCTGGGGCTCAGTGCCCGCTGCTCTGTTGTAACATCTAC 266
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 244 GCCCTGGGCTGAGGCGCGGTGTCCAGCTGCAACATCTTC 285
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
US-09-416-833
Sequence 6, Application US/09416833
Patent No. 6197521
GENERAL INFORMATION:
APPLICANT: Gudkov, Andrei
APPLICANT: Kazarov, Alexander
APPLICANT: Mazo, Ilya
APPLICANT: Roninson, Igor B
TITLE OF INVENTION: Methods for Identifying Genetic
TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
TITLE OF INVENTION: Growth in Cancer Cells
NUMBER OF SEQUENCES: 13

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Allgretti & Witcoff, Ltd.
STREET: 10 S. Wacker Drive, Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/416, 833
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/204, 740
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: No. 6197521nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 93,354-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 285 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-416-833-6

```

Query Match 16.4%; Score 68.4; DB 4; Length 285;
Best Local Similarity 56.8%; Pred. No. 3.4e-13;
Matches 126; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 45 GTGAGAACATCAAGCAGGAGAAATTGGACGCCGGGAGATTGAGATTGCAGACGAC 104
   || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 64 GTGCGGACATCGGACTGGCCCTGGGACGGAAGCCTTGATATAGCTGAGAAATGAG 123
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 105 ATGTCGCTGATTTCACATCAGAAACGTGCTCAGGGGGAGAAAGCCCTGGCTGTGCT 164
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 124 ATGCCAGGTTGATGCGCATGGGGGAGATGTACTCAGCCTCCAGCCACTGAAGGTTGCT 183
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 165 AAATATAGGGCTGTACACACATCACAGCCAGACGGGTGTGATTGAGACACTCTGT 224
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 184 CGCATTCGCTGGCTGCCTGCCGACATGACCGTGGAGACTGCTGTCTCATTTGAGACTCTCGTG 243
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 225 GCCCTGGGGCTCAGTGCCCGCTGGTCTGCTTTACATCTAC 266
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 244 GCCCTGGGTGCTGAGGGCGGTGGTCCAGCTGCACACATCTTC 285
   ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15
PCT-US95-02521-6
; Sequence 6, Application PC/TUS9502521
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Methods for Identifying Genetic
; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
; TITLE OF INVENTION: Growth in Cancer Cells
; NUMBER OF SEQUENCES: 13
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02521

```


GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 02:39:45 ; Search time 83.4703 Seconds
(without alignments)
5018.577 Million cell updates/sec

Title: US-09-782-051-1_COPY_529_945

Perfect score: 417
Sequence: 1 gagagaagcagcaaaccaac.....catgatcctgatgatgagg 417

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 639749 segs, 502280978 residues

Total number of hits satisfying chosen parameters: 1279498

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	417	100.0	2563	10	US-09-782-051-1	Sequence 1, Appl1
2	226.4	54.3	636	10	US-09-879-536-850	Sequence 850, App
3	178	42.7	721	10	US-09-925-300-439	Sequence 439, App
4	137.2	32.9	2200	10	US-09-925-301-217	Sequence 217, App
5	137.2	32.9	2429	12	US-10-044-090-344	Sequence 344, App
6	114.4	27.4	793	10	US-09-966-881-9	Sequence 9, Appl1
7	114.4	27.4	3830	9	US-10-037-598-26	Sequence 26, Appl1
8	114.4	27.4	51309	9	US-09-754-853A-4	Sequence 4, Appl1
9	113	27.1	1396	9	US-09-746-660A-101	Sequence 101, App
10	113	27.1	1422	9	US-09-738-626-836	Sequence 836, App
11	113	27.1	1557	9	US-09-746-660A-97	Sequence 97, Appl1
12	108.2	25.9	389	10	US-09-960-352-10073	Sequence 10073, A
13	106.8	25.6	708	9	US-09-738-626-837	Sequence 837, App
14	97.2	23.3	728	10	US-09-770-149-117	Sequence 117, App
15	94.6	22.7	433	10	US-09-960-352-4421	Sequence 4421, App
16	87.4	21.0	1461	10	US-09-759-990-1	Sequence 1, Appl1
17	72.6	17.4	289	10	US-09-799-946-8	Sequence 8, Appl1
18	68.4	16.4	285	10	US-09-799-946-6	Sequence 6, Appl1
19	52.6	12.6	382	10	US-09-960-352-11499	Sequence 11499, A

C	20	42.8	10.3	110	9	US-09-754-853A-179	Sequence 179, App
C	21	35.6	8.5	99	10	US-09-969-373-209	Sequence 209, App
C	22	33.6	8.1	26006	9	US-10-091-504-1963	Sequence 1963, App
C	23	33.6	8.1	26006	10	US-09-764-869-1963	Sequence 1963, App
C	24	33.6	8.1	26006	10	US-09-764-864-1638	Sequence 1638, App
C	25	33.6	8.1	26013	9	US-10-091-504-1961	Sequence 1961, App
C	26	33.6	8.1	26013	10	US-09-764-869-1961	Sequence 1961, App
C	27	33.6	8.1	26013	10	US-09-764-864-1636	Sequence 1636, App
C	28	33.6	8.1	26018	9	US-10-091-504-1962	Sequence 1962, App
C	29	33.6	8.1	26018	10	US-09-764-869-1962	Sequence 1962, App
C	30	33.6	8.1	26018	10	US-09-764-864-1637	Sequence 1637, App
C	31	32.8	7.9	53522	9	US-09-904-968A-1	Sequence 1, Appl1
C	32	31.6	7.6	8670	10	US-09-964-824A-103	Sequence 103, App
C	33	31.6	7.6	8670	10	US-09-969-708-369	Sequence 369, App
C	34	31.6	7.6	3306	9	US-10-059-962-7	Sequence 7, Appl1
C	35	30.8	7.4	457	10	US-09-770-444-515	Sequence 515, App
C	36	30.6	7.3	594	10	US-09-864-761-7563	Sequence 7563, App
C	37	30.6	7.3	1274	9	US-10-085-188-1	Sequence 1, Appl1
C	38	30.6	7.3	1473	9	US-09-995-898A-3	Sequence 28, Appl1
C	39	30.6	7.3	1560	9	US-09-995-898A-28	Sequence 829, App
C	40	30.4	7.3	3507	10	US-09-834-975-829	Sequence 838, App
C	41	30.4	7.3	3507	10	US-09-834-975-838	Sequence 2, Appl1
C	42	30.2	7.2	1317	9	US-10-121-032-2	Sequence 3457, App
C	43	30	7.2	1878	9	US-09-738-626-3457	Sequence 104, App
C	44	29.8	7.1	750	9	US-10-184-644-104	Sequence 104, App
C	45	29.8	7.1	750	9	US-10-184-634-104	Sequence 104, App

ALIGNMENTS

RESULT 1
US-09-782-051-1
Sequence 1, Application US/09782051
Patent No. US20020035078A1
GENERAL INFORMATION:
APPLICANT: Hart, Derek N J
TITLE OF INVENTION: Enzyme having S-adenosyl-L-homocysteine hydrolase
TITLE OF INVENTION: (AHXY) type activity
FILE REFERENCE: 24305 MRB
CURRENT APPLICATION NUMBER: US/09/782,051
CURRENT FILING DATE: 2001-02-14
PRIOR APPLICATION NUMBER: PCT/NZ97/00133
PRIOR FILING DATE: 1997-10-06
PRIOR APPLICATION NUMBER: NZ 299507
PRIOR FILING DATE: 1996-10-04
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 2563
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (3)..(1847)
OTHER INFORMATION: Open reading frame extends without a stop codon
OTHER INFORMATION: for the full 5' nucleotide sequence. The
OTHER INFORMATION: Initiation codon has yet to be identified.
US-09-782-051-1

Query Match 100.0%; Score 417; DB 10; Length 2563;
Best Local Similarity 100.0%; Pred. No. 5.8e-132;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAGAAGCAGCAACCACTCCAGGCGAGCAATTTCTGTGAAGACATCAAGC 60
DB 529 GAGAGAAGCAGCAACCACTCCAGGCGAGCAATTTCTGTGAAGACATCAAGC 588
QY 61 AGCAGAATTGGACGCCGGGAGATTGAGATTGACAGCAAGACATGCTCTGATT 120
DB 589 AGCAGAATTGGACGCCGGGAGATTGAGATTGACAGCAAGACATGCTCTGATT 648
QY 121 CACTCAGGAACGCTGCTCAGGGGAGAGAGCCCTTGCTGCTAAATAGTGGCTGTA 180

```
|||||
Db 649 CACTCAGGAACGCTGCTCAGGGGAGAGACCCCTTGCTGCTCTAAATAGTGGCTGTA 708
QY 181 CACACATCACAGCCAGACAGCGGTGTGATGAGACACTCTGTGCCCTGGGGCTCAGT 240
Db 709 CACACATCACAGCCAGACAGCGGTGTGATGAGACACTCTGTGCCCTGGGGCTCAGT 768
QY 241 GCCGCTGTCTGCTGTAACTCTACTCACTCAGAATGAGTAGCTGCAGCACTGGCTG 300
Db 769 GCCGCTGTCTGCTGTAACTCTACTCACTCAGAATGAGTAGCTGCAGCACTGGCTG 828
QY 301 AGGCTGAGTTGAGTGTTCGCTTGAAGGGCGAGTCAGAAGATGACTTCGTGGTGTGA 360
Db 829 AGGCTGAGTTGAGTGTTCGCTTGAAGGGCGAGTCAGAAGATGACTTCGTGGTGTGA 888
QY 361 TTGACCGCTGTGTGAACATGATGGGTGGCAGGCCAACATGATCCTGATGATGGG 417
Db 889 TTGACCGCTGTGTGAACATGATGGGTGGCAGGCCAACATGATCCTGATGATGGG 945
```

RESULT 2

```
US-09-879-536-850/c
; Sequence 850, Application US/09879536
; Patent No. US20020144298A1
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dertl, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/879,536
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/088,801
; PRIOR FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 850
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(636)
; OTHER INFORMATION: n - A,T,C or G
US-09-879-536-850
```

```
Query Match 54.3%; Score 226.4; DB 10; Length 636;
Best Local Similarity 99.2%; Pred. No. 3.6e-67;
Matches 238; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
```

```
QY 179 TACACACATCACAGCCAGACAGC-GGTGTGATGAGACACTCTGTGCCCTGGGGCTC 237
Db 420 TACACACATCACAGCCAGACAGC-GGTGTGATGAGACACTCTGTGCCCTGGGGCTC 361
QY 238 AGTCCCGCTGCTGCTGTAACTCTACTCACTCAGAATGAGTAGCTGCAGCACTGG 297
Db 360 AGTCCCGCTGCTGCTGTAACTCTACTCACTCAGAATGAGTAGCTGCAGCACTGG 301
QY 298 CTGAGGCTGAGTTGAGTGTTCGCTTGAAGGGCGAGTCAGAAGATGACTTCGTGGT 357
Db 300 CTGAGGCTGAGTTGAGTGTTCGCTTGAAGGGCGAGTCAGAAGATGACTTCGTGGT 241
QY 358 GTATTGACCGCTGTGTGAACATGATGGGTGGCAGGCCAACATGATCCTGATGATGGG 417
```

```
|||||
Db 240 GTATTGACCGCTGTGTGAACATGATGGGTGGCAGGCCAACATGATCCTGATGATGGG 181
```

RESULT 3

```
US-09-925-300-439
; Sequence 439, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 439
; LENGTH: 721
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (688)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-439
```

```
Query Match 42.7%; Score 178; DB 10; Length 721;
Best Local Similarity 94.2%; Pred. No. 1.3e-50;
Matches 178; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 1 GAGAGAAGCAGCAACCACTCCAGAGGCGAGCAGCAATTTCTGTGAGAACATCAAGC 60
Db 532 GAGAGAAGCAGCAACCACTCCAGAGGCGAGCAGCAATTTCTGTGAGAACATCAAGC 591
QY 61 AGCAGAAATTGGACGCGCGGAGATTGAGATTGCAGAGCAAGACATGTCTGCTGATT 120
Db 592 AGCAGAAATTGGACGCGCGGAGATTGAGATTGCAGAGCAAGACATGTCTGCTGATT 651
QY 121 CACTCAGGAACGCTGCTCAGGGGAGAGAGCCCTTGCTGGTGTCTAAATAGTGGCTGA 180
Db 652 CACTCAGGAACGCTGCTCAGGGGAGAGAGCCCTTGCTGGTGTCTAAATAGTGGCTGA 711
QY 181 CACACATCA 189
Db 712 CACATTTACA 720
```

RESULT 4

```
US-09-925-301-217
; Sequence 217, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 217
; LENGTH: 2200
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```

; NAME/KEY: misc_feature
; LOCATION: (2188)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-217

```

Query Match	32.98;	Score 137.2;	DB 10;	Length 2200;
Best Local Similarity	60.98;	Pred. No. 1.7e-36;		
Matches 223; Conservative	0;	Mismatches 143;	Indels 0;	Gaps 0;

QY	52	ACATCAAGCAGGCAGAAATTGGACGCCGGGAGATTGAGATTGCAGAGCAAGACATGTCTG	111
Db	108	ACATCGGCCTGGCTGCCCTGGGGACGCAAGCCCTTGACATTGCTGAGAACGAGATGCCGG	167
QY	112	CTCTGATTTCACTCAGGAACGTGCTCAGGGGGAGAAGCCCTTGGCTGGTGTCTAAAATAG	171
Db	168	GCCGTGATGCGTATGCGGGAGCGGTACTCGGCCCTCCAAGCCACTGAAGGGCGCCGCATCG	227
QY	172	TGGGCTGTACACACATCACAGCCAGACAGCGGGTGTGATTGAGACACTCTGTGCCCTGG	231
Db	228	CTGGCTGCCCTGCACATGACCCTGGAGACGGCCGTCCTCATTTGAGACCCTGTCACCCCTGG	287
QY	232	GGGCTCAGTGCCGCTGGTCTGCTTGTAACATCTACTCAACTCAGAAATGAAGTAGCTGACG	291
Db	288	GTGCTGAGGTGCAGTGTCTCAGCTGCACAACATCTTCTCCACCAGACCATGGCGGGCTG	347
QY	292	CACCTGGCTGAGGCTGGAGTTGCAGTGTTCGCTTGGAAGGGCGAGTCAGAAAGATGACTTCT	351
Db	348	CCATTGCCAAGGCTGGCATTTCCGGGTATGCTTGGAAGGGCGAAACGAGCAGAGAGTACC	407
QY	352	GGTGGTGTATTGACCGCTGTGTGAACATGGATGGGTGGCAGGCCAACATGATCTCGATG	411
Db	408	TGTGCTGATTTGAGCAGACACCCCTGTACTTCAAGAGACGGGCCCTCAACATGATTTCTGGACG	467
QY	412	ATGGGG 417	
Db	468	ACGGGG 473	

```

RESULT 5
US-10-044-090-344
; Sequence 344, Application US/10044090
; Patent No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044, 090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 344
; LENGTH: 2429
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 1468237CB1
US-10-044-090-344

```

[illegible]

D_b 228 CTGGCTGCCCTGCACATGACCCTGGAGAGCGCCGTCCTCATTTGAGACCCTCGTCAACCCCTGG 287
Q_y 232 GGGCTCAGTGCCGCTGGTCTGTGTTGAACATCTACTCAACTCAGAATGAAGTAGCTGCAG 291
D_b 288 GTGCTGAGGTGCAGTGGTCCAGCTGCACAATCTTCTCACCCAGGACCATGGCGGGCTG 347
Q_y 292 CACTGGCTGAGGCTGGAGTGTGACGTGTTCGCTTGGAAAGGCGGAGTCAGAAGATGACTTCT 351
D_b 348 CCATTGCCAAGGCTGGCATTCGGGTGTATGCCCTGGAAAGGGCGAALCGGACGAGAGTACC 407
Q_y 352 GGTTGGTATTGACCGCTGTGTGAACATGATGGGTGGCAGGCCAACATGATCTTGATG 411
D_b 408 TGTGTGCATTGAGCAGACCCCTGTACTTCAAGGACGGCCCCCTCAACATGATCTGGACG 467
Q_y 412 ATGGCG 417
D_b 468 ACGGGG 473

```

RESULT 6
US-09-966-881-9
; Sequence 9, Application US/09966881
; Patent No. US20020120960A1
; GENERAL INFORMATION:
; APPLICANT: Seymour, Graham
;           Bird, Colin
;           Medina-Suarez, Rosybel
; TITLE OF INVENTION: Genetic control of Fruit Ripening
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zeneca Ag Products Inc.
; STREET: 1800 Concord Pike
; CITY: Wilmington
; STATE: DE
; COUNTRY: USA
; ZIP: 19850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/966,881
; FILING DATE: 28-Sep-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/242,860
; FILING DATE: 29-Mar-1999
; APPLICATION NUMBER: GB 9618862.8
; FILING DATE: 10-SEP-1996
; APPLICATION NUMBER: GB 9708366.1
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: PCT/GB97/02424
; FILING DATE: 08-SEP-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Hohenschutz, Liza D.
; REGISTRATION NUMBER: 33,712
; REFERENCE/DOCKET NUMBER: SEE 50183/UST
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 886-1699
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 793 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: U-0131
; SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-09-966-881-9

```

Query Match	27.48;	Score 114.4;	DB 10;	Length 793;
-------------	--------	--------------	--------	-------------

RESULT 9

US-09-746-660A-101

; Sequence 101, Application US/09746660A
; Publication No. US20030049804A1

; GENERAL INFORMATION:

; APPLICANT: Pompejus, Markus

; APPLICANT: Kroeger, Burkhard

; APPLICANT: Schroder, Hartwig

; APPLICANT: Zelder, Oskar

; APPLICANT: Haberhauer, Gregor

; APPLICANT: Kim, Jun-Won

; APPLICANT: Lee, Heung-Schick

; APPLICANT: Hwang, Byung-Joon

; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING

; FILE REFERENCE: BGI-121CP2

; CURRENT APPLICATION NUMBER: US/09/746,660A

; CURRENT FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 09/606740

; PRIOR FILING DATE: 2000-06-23

; PRIOR APPLICATION NUMBER: 09/603124

; PRIOR FILING DATE: 2000-06-23

; PRIOR APPLICATION NUMBER: 60/141031

; PRIOR FILING DATE: 1999-06-25

; PRIOR APPLICATION NUMBER: 60/142101

; PRIOR FILING DATE: 1999-07-02

; PRIOR APPLICATION NUMBER: 60/148613

; PRIOR FILING DATE: 1999-08-12

; PRIOR APPLICATION NUMBER: 60/187970

; PRIOR FILING DATE: 2000-03-09

; PRIOR APPLICATION NUMBER: DE 19931420.9

; PRIOR FILING DATE: 1999-07-08

; NUMBER OF SEQ ID NOS: 125

; SOFTWARE: PatentIn Vers. 2.0

; SEQ ID NO 101

; LENGTH: 1396

; TYPE: DNA

; ORGANISM: Corynebacterium glutamicum

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (101)..(1396)

; OTHER INFORMATION: FRXA01371

US-09-746-660A-101

Query Match 27.1%; Score 113; DB 9; Length 1396;
Best local similarity 59.4%; Pred. No. 2.5e-28;
Matches 225; Conservative 0; Mismatches 130; Indels 24; Gaps 1;

OY 63 GCAGAAATTGGACGCCGGAGATTGAGATTGCAGAGCAGACATGCTCTGTGATTCA 122
DB 143 GCAGAGGACGAGCAGTCAACGATTCCTTGCAGAGTATGAGATGCCAGGCTCATGAC 202
OY 123 CTGAGGAACGCTGCTCAGGGGAGAGCCCTTGGCTGGTCTAAATAGTGGGCTGTACA 182
DB 203 TTGGCAAGGAATTGCGAGAGCAGCCTTTGAAGGGCGCCGGAATTGCTGTTCTATC 262
OY 183 CACATCACAGCCAGACAGCGGCTTGAATTGAGACACTCTGCTCCCTGGGGGCTCAGTGC 242
DB 263 CACATGACGGTCCAGACCGCGCTTATGAGACCTCACTGCTTTGGCGCTGAGGTT 322
OY 243 CGCTGCTGCTTGTATACATCTACTCACTCAGATGAAGTAGCTGACACTGGCTG-- 300
DB 323 CGTTGGGCTTCTGCAACATTTCTCCACCCAGAGTGAAGGCTGACGGCTATCGTTGC 382
OY 301 -----AGGCTGAGTTCAGTGTTCGCTTGGAGGGGCGAGTCA 338
DB 383 GGCTCCGGCACCCTGGAAGAGCCAGCTGTTCCAGTATTCGCTGGAAGGGTGAAGTCA 442
OY 339 GAAGATGACTTCTGTTGTTATGACCGCTGTGTGAACATGATGGGTGAGGCCAAC 398
DB 443 CTGAGGAGTACTGTTGTTGATCAACAGATCTTCAAGCTGGGGGATGAGCTGCCAAC 502

OY 399 ATGATCCTGATGATGGG 417
DB 503 ATGATCCTGACGACGCGC 521

RESULT 10

US-09-738-626-836

; Sequence 836, Application US/09738626
; Publication No. US20020197605A1

; GENERAL INFORMATION:

; APPLICANT: NAKAGAWA, SATOSHI

; APPLICANT: MIZOGUCHI, HIROSHI

; APPLICANT: ANDO, SEIKO

; APPLICANT: HAYASHI, MIKIRO

; APPLICANT: OCHIAI, KEIKO

; APPLICANT: YOKOI, HARUHIKO

; APPLICANT: TATEISHI, NAOKO

; APPLICANT: SENOH, AKIHIRO

; APPLICANT: IKEDA, MASATO

; APPLICANT: OZAKI, AKIO

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-125

; CURRENT APPLICATION NUMBER: US/09/738,626

; CURRENT FILING DATE: 2000-12-18

; PRIOR APPLICATION NUMBER: JP 99/377484

; PRIOR FILING DATE: 1999-12-16

; PRIOR APPLICATION NUMBER: JP 00/159162

; PRIOR FILING DATE: 2000-04-07

; PRIOR APPLICATION NUMBER: JP 00/280988

; PRIOR FILING DATE: 2000-08-03

; NUMBER OF SEQ ID NOS: 7059

; SOFTWARE: PatentIn ver. 3.0

; SEQ ID NO 836

; LENGTH: 1422

; TYPE: DNA

; ORGANISM: Corynebacterium glutamicum

US-09-738-626-836

Query Match 27.1%; Score 113; DB 9; Length 1422;
Best local similarity 59.4%; Pred. No. 2.5e-28;
Matches 225; Conservative 0; Mismatches 130; Indels 24; Gaps 1;

OY 63 GCAGAAATTGGACGCCGGAGATTGAGATTGCAGAGCAGACATGCTCTGTGATTCA 122
DB 31 GCAGAGGACGAGCAGTCAACGATTCCTTGCAGAGTATGAGATGCCAGGCTCATGAC 90
OY 123 CTGAGGAACGCTGCTCAGGGGAGAGCCCTTGGCTGGTCTAAATAGTGGGCTGTACA 182
DB 91 TTGGCAAGGAATTGCGAGAGCAGCCTTTGAAGGGCGCCGGAATTGCTGTTCTATC 150
OY 183 CACATCACAGCCAGACAGCGGCTTGAATTGAGACACTCTGCTCCCTGGGGGCTCAGTGC 242
DB 151 CACATGACGGTCCAGACCGCGCTTATGAGACCTCACTGCTTTGGCGCTGAGGTT 210
OY 243 CGCTGCTGCTTGTATACATCTACTCACTCAGATGAAGTAGCTGACACTGGCTG-- 300
DB 211 CGTTGGGCTTCTGCAACATTTCTCCACCCAGAGTGAAGGCTGACGGCTATCGTTGC 270
OY 301 -----AGGCTGAGTTCAGTGTTCGCTTGGAGGGGCGAGTCA 338
DB 271 GGCTCCGGCACCCTGGAAGAGCCAGCTGTTCCAGTATTCGCTGGAAGGGTGAAGTCA 330
OY 339 GAAGATGACTTCTGTTGTTATGACCGCTGTGTGAACATGATGGGTGAGGCCAAC 398
DB 331 CTGAGGAGTACTGTTGTTGATCAACAGATCTTCAAGCTGGGGGATGAGCTGCCAAC 390
OY 399 ATGATCCTGATGATGGG 417
DB 391 ATGATCCTGACGACGCGC 409

RESULT 11

US-09-746-660A-97
; Sequence 97, Application US/09746660A
; Publication No. US20030049804A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Krogger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; APPLICANT: Kim, Jun-Won
; APPLICANT: Lee, Heung-Schick
; APPLICANT: Hwang, Byung-Joon
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
; TITLE OF INVENTION: METABOLIC PATHWAY PROTEINS
; FILE REFERENCE: BGI-121CP2
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US/09/746,660A
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 09/603124
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142101
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/148613
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/187970
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn Vers. 2.0
; SEQ ID NO 97
; LENGTH: 1557
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(1534)
; OTHER INFORMATION: RXN00132
US-09-746-660A-97

Query Match 27.1%; Score 113; DB 9; Length 1557;
Best Local Similarity 59.4%; Pred. No. 2.6e-28;
Matches 225; Conservative 0; Mismatches 130; Indels 24; Gaps 1;

QY 63 GCAGATTGGACGCCGGAGATTGAGATTGCAGACAGACATGCTGCTCTGATTCA 122
DB 143 GCAGAGGACGACGTCACAGATTGCTTGCAGATATGAGATGCCAGGTCATGCAG 202
QY 123 CTCAGGAACGTCCTCAGGGGAGAGAGCCCTTGCTGCTAAATAGTGGCTGTACA 182
DB 203 TTGGCAAGGAATTCGACAGACAGACGCTTGAAGGGCCCGCAATTGCTGTTCTATC 262
QY 183 CACATCACAGCCACAGACGCGTGTGATTGAGACACTCTGTGCCCTGGGGCTCAGTGC 242
DB 263 CACATGACGGTCCAGACCGCGCTGCTTATGAGACCTCACTGCTTGGGGCTGAGGTT 322
QY 243 CGCTGCTCTGCTTGTACATCTACTCACTCAGATGAAGTAGCTGCAGCAGCTGGCTG-- 300
DB 323 CGTGGGCTTCTGCAACATTTCTCCACCCAGAGTAGGGCTGCAGCGGCTATGTTGTC 382
QY 301 -----AGGCTGAGTTGCAAGTGTGCTTGGTGAAGGGCGAGTCA 338
DB 383 GGCTCGGACCGTCGAGAGAGCCAGCTGCTGCTCCAGTATTCGCGTGAAGGGTGAAGTCA 442
QY 339 GAAGATGACTTCTGTGGTGTATTGACCGCTGTGTGAACATGATGGGTGCAGGCCAAC 398
DB 443 CTGAGAGAGTACTGCTGTCATCAACAGATCTTCAGCTGGGGCGATGAGCTGCCAAC 502
QY 399 ATGATCCTGATGATGGG 417
|||||

DB 503 ATGATCCTGACGACGGCG 521

RESULT 12
US-09-960-352-10073
; Sequence 10073, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathalagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 10073
; LENGTH: 389
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (284)..(292)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: 43-LIB34-036-Q1-E1-C4
US-09-960-352-10073

Query Match 25.9%; Score 108.2; DB 10; Length 389;
Best Local Similarity 61.1%; Pred. No. 6.4e-27;
Matches 173; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 52 ACATCAAGCAGCGAGAAATTGGACGCCGGGAGATTGAGATTGCAGAGCAAGACATGCTG 111
DB 105 ACATCAGCCTGGCCGCTTGGGGACGCAAGGCCCTGGACCTGGCAGAGATGATGCCG 164
QY 112 CTCTGATTTCACCTCAGGAACGCTGCTCAGGGGAGAGCCCTGGCTGCTGCTAAATAG 171
DB 165 GCCTGATGACATGCGAGAAATGTAAGCTGCGCTCAAGCCTGGAAGGGCGCTGCAATG 224
QY 172 TGGGCTGTACACATCAACAGCCAGACAGCGGTTGATTGAGACACTCTGTCCTGG 231
DB 225 CTGGCTGCTGACATGACATGACCGTGGAGACCGCCGCTCCATGAGACCTGTCCTG 284
QY 232 GGGCTCAGTCCCGCTGCTGCTGTTGATACATCTACTCACTCAGATGAGTAGCTGCA 291
DB 285 GTGCTGANGTGGGCTGCTCAGCTGCAATATCTTCTCCACCCAGGACCATGACGACTG 344
QY 292 CACTGCTGAGGCTGAGTTCAGTGTGCTGCTGGAAGGGCGA 334
DB 345 CCATTGCCAAGGCTGCAATTCAGTGTACGCGCTGGAAGGGTGA 387

RESULT 13
US-09-738-626-837/c
; Sequence 837, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHITO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT FILING DATE: 2000-12-18

Db	97	CAGGACCATGACAGCAGCTGCCAATGCCAAGGCTGGCATCCAGTGACGCCCTGGAAGGT	156
QY	333	GAGTCAGAAGATGACTTCTGGTGGTGTATTGACCCGCTGTGTGAACATGATGGGTGGCAG	392
Db	157	GAACGGATGAGAGTACTCTGTGTCATTGAACAGACGCGTGTACTTCAAGGACGGGCC	216
QY	393	GCCACATGATCCTGATGATGGG	417
Db	217	CTCACATGATTCTGACGACGCTG	241

Search completed: April 21, 2003, 07:21:15
Job time : 336.47 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2003, 23:31:25 ; Search time 562.956 Seconds
(without alignments)
11996.532 Million cell updates/sec

Title: US-09-782-051-1_COPY_529_945

Perfect score: 417
Sequence: 1 gagagagacagcaaccaac.....catgatcctgatgatgaggg 417

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	417	100.0	700	9	AU133884	AU133884 AU133884
2	417	100.0	731	12	BG826655	BG826655 602748977
3	417	100.0	744	12	BG699160	BG699160 602678818
4	417	100.0	763	12	BG706498	BG706498 602670053
5	417	100.0	764	9	AU126038	AU126038 AU126038
6	417	100.0	791	12	BG701625	BG701625 602682515

7	417	100.0	867	13	BM453920	BM453920 AGENCOURT
8	417	100.0	921	9	AL551097	AL551097 AL551097
9	417	100.0	944	12	BF033444	BF033444 601457970
10	417	100.0	969	12	BG259448	BG259448 602378544
11	416.6	99.9	858	9	AL554850	AL554850 AL554850
12	416.6	99.9	938	9	AL527928	AL527928 AL527928
13	415.4	99.6	701	12	BG706693	BG706693 602671806
14	415.4	99.6	857	13	BI668412	BI668412 603292688
15	414	99.3	1042	13	BM449470	BM449470 AGENCOURT
16	406	97.4	773	13	BI549142	BI549142 603189237
17	404.4	97.0	822	13	BI460083	BI460083 603301645
18	395.8	94.9	964	10	AV723709	AV723709 AV723709
19	389.2	93.3	1088	12	BF797186	BF797186 602257870
20	388.6	93.2	632	12	BF307988	BF307988 601894234
21	384.2	92.1	929	12	BE781733	BE781733 601467403
22	383.4	91.9	529	9	AI391414	AI391414 mw74d10.y
23	383.4	91.9	785	12	BG699237	BG699237 602678908
24	382.8	91.8	712	13	BI764959	BI764959 603047381
25	375.4	90.0	553	10	BE279230	BE279230 601156734
26	374	89.7	589	14	BM750989	BM750989 K-EST0026
27	373	89.4	665	12	BF002488	BF002488 7h07e06.x
28	371.6	89.1	680	10	AV724011	AV724011 AV724011
29	370	88.7	937	13	BI597110	BI597110 603251755
30	369.2	88.5	769	13	BI551275	BI551275 603194678
31	368.4	88.3	478	12	BG314811	BG314811 OP2.0.231
32	363.4	87.1	701	10	BE278319	BE278319 601179921
33	362	86.8	661	13	BI544920	BI544920 603242371
34	359.4	86.2	656	13	BI596544	BI596544 603243595
35	352.2	84.5	724	9	AU080435	AU080435 AU080435
36	352	84.4	747	12	BG697210	BG697210 602660444
37	350.8	84.1	818	12	BG400835	BG400835 602464057
38	349	83.7	891	9	AL520704	AL520704 AL520704
39	348	83.5	963	12	BF530850	BF530850 602072525
40	340	81.5	865	9	AL519571	AL519571 AL519571
41	339.4	81.4	768	12	BE914663	BE914663 601665173
42	337	80.8	748	9	AU124992	AU124992 AU124992
43	331.2	79.4	750	9	AJ453210	AJ453210 AJ453210
44	328.8	78.8	570	12	BG482072	BG482072 602527956
45	324.2	77.7	687	10	BE276116	BE276116 601121486

ALIGNMENTS

RESULT 1
LOCUS AU133884 700 bp mRNA linear EST 01-AUG-2002
DEFINITION AU133884 OVARC1 Homo sapiens cDNA clone OVARC100852 5', mRNA
ACCESSION AU133884
VERSION AU133884.1 GI:10994423
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 700)
AUTHORS Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,T.
TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix
Research Institute; cDNA library construction; Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.

FEATURES
Source Location/Qualifiers
1. 700
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="OVARC1000852"
/tissue_type="OVARC1"
/tissue_type="ovary, tumor tissue"
/note="Vector: pME18SFL3"

BASE COUNT 178 a 154 c 197 g 168 t 3 others

ORIGIN

Query Match 100.0%; Score 417; DB 9; Length 700;
Best Local Similarity 100.0%; Pred. No. 2.3e-110;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAGAAGCAGCAACCACTCCAAAGGCGAGCAGCAATTTCTGTGAGAACAATCAAC 60
|||
Db 93 GAGAGAAGCAGCAACCACTCCAAAGGCGAGCAGCAATTTCTGTGAGAACAATCAAC 152
|||
QY 61 AGCAGAAATTTGGACGCGGGAGATTGAGATTGACAGACAGACATGTCTGCTGATTT 120
|||
Db 153 AGCAGAAATTTGGACGCGGGAGATTGAGATTGACAGACAGACATGTCTGCTGATTT 212
|||
QY 121 CACTCAGAAACGTGCTCAGGCGGAGAAAGCCCTTGGCTGCTAAATAGTGGCTGTA 180
|||
Db 213 CACTCAGAAACGTGCTCAGGCGGAGAAAGCCCTTGGCTGCTAAATAGTGGCTGTA 272
|||
QY 181 CACACATCACAGCCAGACAGCGGTGTTGATTGAGACACTCTGTGCCCTGGGGCTCAGT 240
|||
Db 273 CACACATCACAGCCAGACAGCGGTGTTGATTGAGACACTCTGTGCCCTGGGGCTCAGT 332
|||
QY 241 GCCGCTGCTGCTGTGTAACATCTACTCACTCAGAAATGAGTAGCTGCAGCAGTGGCTG 300
|||
Db 333 GCCGCTGCTGCTGTGTAACATCTACTCACTCAGAAATGAGTAGCTGCAGCAGTGGCTG 392
|||
QY 301 AGCCTGAGTTGCAAGTGTTCCTTGGAAAGGCGGAGTCAAGATGACTTCTGCTGCTGA 360
|||
Db 393 AGCCTGAGTTGCAAGTGTTCCTTGGAAAGGCGGAGTCAAGATGACTTCTGCTGCTGA 452
|||
QY 361 TTGACCGCTGTGTGAACATGATGGGTGGCAGGCCCAACATGATCCTGGATGATGGGG 417
|||
Db 453 TTGACCGCTGTGTGAACATGATGGGTGGCAGGCCCAACATGATCCTGGATGATGGGG 509
|||

RESULT 2
BG826655 731 bp mRNA linear EST 22-MAY-2001
LOCUS 602748977F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4901851 5',
DEFINITION mRNA sequence.
ACCESSION BG826655
VERSION BG826655.1 GI:14174242
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 731)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LNCMI797 row: 1 column: 20
High quality sequence stop: 728.
Location/Qualifiers
1. 731

FEATURES
Source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4901851"
/clone_1b="NIH_MGC_17"
/tissue_type="rhabdomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI;
Site_2: XhoI; cDNA made by oligo-dT priming.
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 185 a 161 c 210 g 173 t 2 others

ORIGIN

Query Match 100.0%; Score 417; DB 12; Length 731;
Best Local Similarity 100.0%; Pred. No. 2.4e-110;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAGAAGCAGCAACCACTCCAAAGGCGAGCAGCAATTTCTGTGAGAACAATCAAC 60
|||
Db 96 GAGAGAAGCAGCAACCACTCCAAAGGCGAGCAGCAATTTCTGTGAGAACAATCAAC 155
|||
QY 61 AGCAGAAATTTGGACGCGGGAGATTGAGATTGACAGACAGACATGTCTGCTGATTT 120
|||
Db 156 AGCAGAAATTTGGACGCGGGAGATTGAGATTGACAGACAGACATGTCTGCTGATTT 215
|||
QY 121 CACTCAGAAACGTGCTCAGGCGGAGAAAGCCCTTGGCTGCTAAATAGTGGCTGTA 180
|||
Db 216 CACTCAGAAACGTGCTCAGGCGGAGAAAGCCCTTGGCTGCTAAATAGTGGCTGTA 275
|||
QY 181 CACACATCACAGCCAGACAGCGGTGTTGATTGAGACACTCTGTGCCCTGGGGCTCAGT 240
|||
Db 276 CACACATCACAGCCAGACAGCGGTGTTGATTGAGACACTCTGTGCCCTGGGGCTCAGT 335
|||
QY 241 GCCGCTGCTGCTGTGTAACATCTACTCACTCAGAAATGAGTAGCTGCAGCAGTGGCTG 300
|||
Db 336 GCCGCTGCTGCTGTGTAACATCTACTCACTCAGAAATGAGTAGCTGCAGCAGTGGCTG 395
|||
QY 301 AGCCTGAGTTGCAAGTGTTCCTTGGAAAGGCGGAGTCAAGATGACTTCTGCTGCTGA 360
|||
Db 396 AGCCTGAGTTGCAAGTGTTCCTTGGAAAGGCGGAGTCAAGATGACTTCTGCTGCTGA 455
|||
QY 361 TTGACCGCTGTGTGAACATGATGGGTGGCAGGCCCAACATGATCCTGGATGATGGGG 417
|||
Db 456 TTGACCGCTGTGTGAACATGATGGGTGGCAGGCCCAACATGATCCTGGATGATGGGG 512
|||

RESULT 3
BG699160 744 bp mRNA linear EST 07-MAY-2001
LOCUS 602678818F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4811477 5',
DEFINITION mRNA sequence.
ACCESSION BG699160
VERSION BG699160.1 GI:13967178
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 744)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1AM10702 row: d column: 06
High quality sequence stop: 741.

FEATURES

source

1. 744

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4811477"

/clone_lib="NIH_MGC_95"

/tissue_type="hippocampus"

/lab_host="DH10B"

/note="Organ: brain; Vector: pBluescript (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3', size-selected for average insert size 2.5 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH-MGC Library."

BASE COUNT 194 a 175 c 217 g 158 t
ORIGIN

Query Match

Best Local Similarity 100.0%; Score 417; DB 12; Length 744;

Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAGAAGCAGCAACCACTCCAAGGCGAGCATTTCTGTGAGAGAACATCAAGC 60

Db 241 GAGAGAAGCAGCAACCACTCCAAGGCGAGCATTTCTGTGAGAGAACATCAAGC 300

QY 61 AGGCAGATTGTGACGCCGGGAGATTGAGATTGCAGAGACATGTCTGCTGATTT 120

Db 301 AGGCAGATTGTGACGCCGGGAGATTGAGATTGCAGAGACATGTCTGCTGATTT 360

QY 121 CACTCAGGAACGTCGTGAGGGGAGAGAGCCCTGGCTGCTAAATAGTGGCTGTA 180

Db 361 CACTCAGGAACGTCGTGAGGGGAGAGAGCCCTGGCTGCTAAATAGTGGCTGTA 420

QY 181 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTCTGCCCCGGGGCTCAGT 240

Db 421 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTCTGCCCCGGGGCTCAGT 480

QY 241 GCCGCTGCTGCTGTGTAACATCTACTCAACTCAGATGAGTAGCTGCAGCACTGGCTG 300

Db 481 GCCGCTGCTGCTGTGTAACATCTACTCAACTCAGATGAGTAGCTGCAGCACTGGCTG 540

QY 301 AGCCTGAGTTGAGTGTGCTTGAAGGGGAGTCAAGATGACTTCTGTGTGTGTA 360

Db 541 AGCCTGAGTTGAGTGTGCTTGAAGGGGAGTCAAGATGACTTCTGTGTGTGTA 600

QY 361 TTGACCGCTGTGTGAACATGATGGGTGCAGGCCAATGATCCTGATGATGGG 417

Db 601 TTGACCGCTGTGTGAACATGATGGGTGCAGGCCAATGATCCTGATGATGGG 657

RESULT 4 763 bp mRNA linear EST 07-MAY-2001

LOCUS BG706498 602670053F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4792813 5'

DEFINITION mRNA sequence.

ACCESSION BG706498

VERSION BG706498.1 GI:13981903

KEYWORDS

SOURCE

ORGANISM human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 763)

NIH-MGC http://mgc.ncl.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL COMMENT

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki

Toshiyuki and Piero Carninci (RIKEN)

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1AM10671 row: j column: 14

High quality sequence stop: 737.

FEATURES

source

1. 763

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4792813"

/clone_lib="NIH_MGC_96"

/tissue_type="hypothalamus"

/lab_host="DH10B"

/note="Organ: brain; Vector: pBluescript (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH-MGC Library."

BASE COUNT 195 a 176 c 224 g 168 t
ORIGIN

Query Match

Best Local Similarity 100.0%; Score 417; DB 12; Length 763;

Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAGAAGCAGCAACCACTCCAAGGCGAGCATTTCTGTGAGAGAACATCAAGC 60

Db 295 GAGAGAAGCAGCAACCACTCCAAGGCGAGCATTTCTGTGAGAGAACATCAAGC 354

QY 61 AGGCAGATTGTGACGCCGGGAGATTGAGATTGCAGAGCAATGCTGCTGATTT 120

Db 355 AGGCAGATTGTGACGCCGGGAGATTGAGATTGCAGAGCAATGCTGCTGATTT 414

QY 121 CACTCAGGAACGTCGTGAGGGGAGAGAGCCCTGGCTGCTAAATAGTGGCTGTA 180

Db 415 CACTCAGGAACGTCGTGAGGGGAGAGAGCCCTGGCTGCTAAATAGTGGCTGTA 474

QY 181 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTCTGCCCCGGGGCTCAGT 240

Db 475 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTCTGCCCCGGGGCTCAGT 534

QY 241 GCCGCTGCTGCTGTGTAACATCTACTCAACTCAGATGAGTAGCTGCAGCACTGGCTG 300

Db 535 GCCGCTGCTGCTGTGTAACATCTACTCAACTCAGATGAGTAGCTGCAGCACTGGCTG 594

QY 301 AGCCTGAGTTGAGTGTGCTTGAAGGGGAGTCAAGATGACTTCTGTGTGTGTA 360

Db 595 AGCCTGAGTTGAGTGTGCTTGAAGGGGAGTCAAGATGACTTCTGTGTGTGTA 654

QY 361 TTGACCGCTGTGTGAACATGATGGGTGCAGGCCAATGATCCTGATGATGGG 417

Db 655 TTGACCGCTGTGTGAACATGATGGGTGCAGGCCAATGATCCTGATGATGGG 711

RESULT 5

AUI26038

LOCUS AUI26038 764 bp mRNA linear EST 01-AUG-2002

DEFINITION AUI26038 NT2RM4 Homo sapiens cDNA clone NT2RM4002613 5', mRNA

sequence.

ACCESSION AUI26038

VERSION AUI26038.1 GI:10950754

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S., Salto,K., Yamamoto,J.,
Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S. and
Isogai,T.
TITLE HRI human cDNA project (Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S.,
Salto,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki
Y., Sugano,S., Isogai,T.)
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix
Research Institute; cDNA library construction; Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
location/Qualifiers
1. 764
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RM4002613"
/clone_1lb="NT2RM4"
/cell_type="teratocarcinoma"
/note="Vector: pME18SFL3; mRNA from uninduced NT2 neuronal
precursor cells"
BASE COUNT 193 a 157 c 228 g 183 t 3 others
ORIGIN
Query Match 100.0%; Score 417; DB 9; Length 764;
Best Local Similarity 100.0%; Pred. No. 2.4e-110;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGAGAGCAGCAAAACCACTCCAGAGGGCAGCAGCAATTTCTGTGTAAGAACAATCAAGC 60
DB 58 GAGAGAGCAGCAAAACCACTCCAGAGGGCAGCAGCAATTTCTGTGTAAGAACAATCAAGC 117
QY 61 AGCAGAAATTTGGACGCCGGGAGATTGAGATTGACAGAGCAAGACATGCTGCTGATTT 120
DB 118 AGCAGAAATTTGGACGCCGGGAGATTGAGATTGACAGAGCAAGACATGCTGCTGATTT 177
QY 121 CACTCAGGAACGTGCTCAGGGGGAGAAAGCCCTGGCTGCTAATAATAGTGGCTGTA 180
DB 178 CACTCAGGAACGTGCTCAGGGGGAGAAAGCCCTGGCTGCTAATAATAGTGGCTGTA 237
QY 181 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTCTGCCCCCTGGGGCTCAGT 240
DB 238 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTCTGCCCCCTGGGGCTCAGT 297
QY 241 GCCGCTGCTGCTGTGTAACATCTACTCAACTCAGAATGAAGTAGCTGCAGCACTGGCTG 300
DB 298 GCCGCTGCTGCTGTGTAACATCTACTCAACTCAGAATGAAGTAGCTGCAGCACTGGCTG 357
QY 301 AGGCTGAGTGTGACAGTGTGCTGGAAGGGCGAGTCAGAGATGACTTCTGTGTGTA 360
DB 358 AGGCTGAGTGTGACAGTGTGCTGGAAGGGCGAGTCAGAGATGACTTCTGTGTGTA 417
QY 361 TTGACCGCTGTGTGACATGATGGTGGCAGGCCACATGATCTGATGATGGG 417
DB 418 TTGACCGCTGTGTGACATGATGGTGGCAGGCCACATGATCTGATGATGGG 474
RESULT 6
BG701625
LOCUS BG701625 791 bp mRNA linear EST 07-MAY-2001

DEFINITION 602682515F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4815170 5',
mRNA sequence.
ACCESSION BG701625
VERSION BG701625.1 GI:13972152
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Miklos Falkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
Toshiyuki and Ptero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
plate: LLM10711 row: n column: 03
High quality sequence stop: 707.
location/Qualifiers
1. 791
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4815170"
/clone_1lb="NIH_MGC_95"
/tissue_type="hippocampus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.5 kb and
normalized to R0.5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 196 a 193 c 226 g 176 t
ORIGIN
Query Match 100.0%; Score 417; DB 12; Length 791;
Best Local Similarity 100.0%; Pred. No. 2.5e-110;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGAGAGCAGCAAAACCACTCCAGAGGGCAGCAGCAATTTCTGTGTAAGAACAATCAAGC 60
DB 252 GAGAGAGCAGCAAAACCACTCCAGAGGGCAGCAGCAATTTCTGTGTAAGAACAATCAAGC 311
QY 61 AGCAGAAATTTGGACGCCGGGAGATTGAGATTGACAGAGCAAGACATGCTGCTGATTT 120
DB 312 AGCAGAAATTTGGACGCCGGGAGATTGAGATTGACAGAGCAAGACATGCTGCTGATTT 371
QY 121 CACTCAGGAACGTGCTCAGGGGGAGAAAGCCCTGGCTGCTAATAATAGTGGCTGTA 180
DB 372 CACTCAGGAACGTGCTCAGGGGGAGAAAGCCCTGGCTGCTAATAATAGTGGCTGTA 431
QY 181 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTCTGCCCCCTGGGGCTCAGT 240
DB 432 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTCTGCCCCCTGGGGCTCAGT 491
QY 492 GCCGCTGCTGCTGTGTAACATCTACTCAACTCAGAATGAAGTAGCTGCAGCACTGGCTG 551
DB 552 AGGCTGAGTGTGACAGTGTGCTGGAAGGGCGAGTCAGAGATGACTTCTGTGTGTA 611

QY 361 TTGACCGCTGTGTGACATGATGGTGGCAGGCCAACATGATCTGTGATGATGGG 417
|||||
DB 612 TTGACCGCTGTGTGAACATGATGGTGGCAGGCCAACATGATCTGTGATGATGGG 668

RESULT 7
LOCUS BM453920 867 bp mRNA linear EST_05-FEB-2002
DEFINITION AGENCOURT_6402958 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5528995
5', mRNA sequence.
ACCESSION BM453920
VERSION BM453920.1 GI:18502949
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 867)
NIH-MGC <http://mhc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LLM12206 row: 1 column: 20
High quality sequence stop: 670.

FEATURES
Source
1. 867
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5528995"
/clone_1lb="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

BASE COUNT 222 a 201 c 241 g 202 t 1 others
ORIGIN

Query Match 100.0%; Score 417; DB 13; Length 867;
Best Local Similarity 100.0%; Pred. No. 2.6e-110;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAGAGCAGCAAACTCCAGGCGCAGCAGCAATTTCTGTGAGAACATCAAGC 60
|||||
DB 202 GAGAGAGCAGCAAACTCCAGGCGCAGCAGCAATTTCTGTGAGAACATCAAGC 261

QY 61 AGGAGAAATTGGACGCCGGAGATTGACATGACAGACAGACATGCTCTGATTT 120
|||||
DB 262 AGGAGAAATTGGACGCCGGAGATTGACATGACAGACAGACATGCTCTGATTT 321

QY 121 CACTCAGGAAACGTGCTCAGGGGAGAAAGCCCTGGCTGCTAAATAGTGGGCTGTA 180
|||||
DB 322 CACTCAGGAAACGTGCTCAGGGGAGAAAGCCCTGGCTGCTAAATAGTGGGCTGTA 381

QY 181 CACACATCAGAGCCAGACAGCGGCTGTGATGAGACACTCTGTGCCCTGGGGGCTCAGT 240
|||||
DB 382 CACACATCAGAGCCAGACAGCGGCTGTGATGAGACACTCTGTGCCCTGGGGGCTCAGT 441

QY 241 GCCGCTGCTGCTGTGTAACATCTACTCACTCAGATGAAGTAGCTGACAGACTGGCTG 300
|||||
DB 442 GCCGCTGCTGCTGTGTAACATCTACTCACTCAGATGAAGTAGCTGACAGACTGGCTG 501

QY 301 AGGCTGAGTTGCAAGTGTGCTTGGAGGGCGAGTCAGAGATGACTTCTGTGCTGTA 360
|||||

DB 502 AGGCTGAGTTGCAAGTGTGCTTGGAGGGCGAGTCAGAGATGACTTCTGTGCTGTA 561

QY 361 TTGACCGCTGTGTGAACATGATGGTGGCAGGCCAACATGATCTGTGATGATGGG 417
|||||
DB 562 TTGACCGCTGTGTGAACATGATGGTGGCAGGCCAACATGATCTGTGATGATGGG 618

RESULT 8
LOCUS AL551097 921 bp mRNA linear EST_16-FEB-2001
DEFINITION AL551097 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0D1066YM07.5
prime, mRNA sequence.
ACCESSION AL551097
VERSION AL551097.1 GI:12888715
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 921)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
Source
1. 921
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1066YM07"
/clone_1lb="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com>"

BASE COUNT 217 a 231 c 302 g 170 t 1 others
ORIGIN

Query Match 100.0%; Score 417; DB 9; Length 921;
Best Local Similarity 100.0%; Pred. No. 2.7e-110;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAGAGCAGCAAACTCCAGGCGCAGCAGCAATTTCTGTGAGAACATCAAGC 60
|||||
DB 461 GAGAGAGCAGCAAACTCCAGGCGCAGCAGCAATTTCTGTGAGAACATCAAGC 520

QY 61 AGGAGAAATTGGACGCCGGAGATTGACATGACAGACAGACATGCTCTGATTT 120
|||||
DB 521 AGGAGAAATTGGACGCCGGAGATTGACATGACAGACAGACATGCTCTGATTT 580

QY 121 CACTCAGGAAACGTGCTCAGGGGAGAAAGCCCTGGCTGCTAAATAGTGGGCTGTA 180
|||||
DB 581 CACTCAGGAAACGTGCTCAGGGGAGAAAGCCCTGGCTGCTAAATAGTGGGCTGTA 640

QY 181 CACACATCAGAGCCAGACAGCGGCTGTGATGAGACACTCTGTGCCCTGGGGGCTCAGT 240
|||||
DB 641 CACACATCAGAGCCAGACAGCGGCTGTGATGAGACACTCTGTGCCCTGGGGGCTCAGT 700

QY 241 GCCGCTGCTGCTGTGTAACATCTACTCACTCAGATGAAGTAGCTGACAGACTGGCTG 300
|||||
DB 701 GCCGCTGCTGCTGTGTAACATCTACTCACTCAGATGAAGTAGCTGACAGACTGGCTG 760

QY 301 AGGCTGAGTTGCAAGTGTGCTTGGAGGGCGAGTCAGAGATGACTTCTGTGCTGTA 360
|||||

Db 761 AGGCTGAGATTGACGTGTTGCGCTTGAAGGGCGAGTCAGAGATGACTTCTGCTGTA 820
OY 361 TTGACCGCTGTGTGACATGATGGTGGCAGGCCAACATGATCTGATGATGGG 417
|||||
Db 821 TTGACCGCTGTGTGACATGATGGTGGCAGGCCAACATGATCTGATGATGGG 877

RESULT 9
BF033444

LOCUS 944 bp mRNA linear EST:20-OCT-2000
DEFINITION 601457970F1 NIH_MGC_66 Homo sapiens cDNA clone IMAGE:3861641 5',
mRNA sequence.
ACCESSION BF033444
VERSION BF033444.1 GI:10741156
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE NIH-MGC http://mgs.nci.nih.gov/
1 (bases 1 to 944)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM9598 row: k column: 18
High quality sequence stop: 600.

FEATURES

Source

1.944
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3861641"
/clone_lib="NIH_MGC_66"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.8 kb. Library constructed by Life
Technologies."
BASE COUNT 244 a 198 c 273 g 229 t
ORIGIN

Query Match 100.0%; Score 417; DB 12; Length 944;
Best Local Similarity 100.0%; Pred. No. 2.7e-110;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAGAGAAGCAGCAAACTCCAAAGGCGAGCAATTTCTGTGTAAGAACAATCAAGC 60
|||||
Db 89 GAGAGAAGCAGCAAACTCCAAAGGCGAGCAATTTCTGTGTAAGAACAATCAAGC 148
OY 61 AGCAGAAATTGGACGCCGGAGATTGAGATTGACAGACAAGACATGCTGCTGATT 120
|||||
Db 149 AGCAGAAATTGGACGCCGGAGATTGAGATTGACAGACAAGACATGCTGCTGATT 208
OY 121 CACTCAGGAACGCTCTCAGGGGAGAGAAGCCCTTGCTGCTAAATAGTGGGCTGTA 180
|||||
Db 209 CACTCAGGAACGCTCTCAGGGGAGAGAAGCCCTTGCTGCTAAATAGTGGGCTGTA 268
OY 181 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTGTGCCCCGCGGCTCAGT 240
|||||
Db 269 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTGTGCCCCGCGGCTCAGT 328
OY 241 GCCGCTGTCTGTTGTAACATCTACTCAACTCAGAAATGAGTAGCTGCAGCACTGCTG 300
|||||
Db 329 GCCGCTGTCTGTTGTAACATCTACTCAACTCAGAAATGAGTAGCTGCAGCACTGCTG 388

OY 301 AGGCTGAGTTGACGTGTTGCGCTTGAAGGGCGAGTCAGAGATGACTTCTGCTGTA 360
|||||
Db 389 AGGCTGAGTTGACGTGTTGCGCTTGAAGGGCGAGTCAGAGATGACTTCTGCTGTA 448
OY 361 TTGACCGCTGTGTGAACATGATGGTGGCAGGCCAACATGATCTGATGATGGG 417
|||||
Db 449 TTGACCGCTGTGTGAACATGATGGTGGCAGGCCAACATGATCTGATGATGGG 505

RESULT 10
BG259448

LOCUS 969 bp mRNA linear EST:13-FEB-2001
DEFINITION 602378544F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4509347 5',
mRNA sequence.
ACCESSION BG259448
VERSION BG259448.1 GI:12769264
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE NIH-MGC http://mgs.nci.nih.gov/
1 (bases 1 to 969)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM10389 row: g column: 12
High quality sequence stop: 700.

FEATURES

Source

1.969
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4509347"
/clone_lib="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; Oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 234 a 219 c 290 g 226 t
ORIGIN

Query Match 100.0%; Score 417; DB 12; Length 969;
Best Local Similarity 100.0%; Pred. No. 2.7e-110;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAGAGAAGCAGCAAACTCCAAAGGCGAGCAATTTCTGTGTAAGAACAATCAAGC 60
|||||
Db 85 GAGAGAAGCAGCAAACTCCAAAGGCGAGCAATTTCTGTGTAAGAACAATCAAGC 144
OY 61 AGCAGAAATTGGACGCCGGAGATTGAGATTGACAGACAAGACATGCTGCTGATT 120
|||||
Db 145 AGCAGAAATTGGACGCCGGAGATTGAGATTGACAGACAAGACATGCTGCTGATT 204
OY 121 CACTCAGGAACGCTCTCAGGGGAGAGAAGCCCTTGCTGCTAAATAGTGGGCTGTA 180
|||||
Db 205 CACTCAGGAACGCTCTCAGGGGAGAGAAGCCCTTGCTGCTAAATAGTGGGCTGTA 264
OY 181 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTGTGCCCCGCGGCTCAGT 240
|||||
Db 265 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTGTGCCCCGCGGCTCAGT 324
OY 241 GCCGCTGTCTGTTGTAACATCTACTCAACTCAGAAATGAGTAGCTGCAGCACTGCTG 300

```
|||||
Db 325 GCCGCTGCTCTGCTGTGTACATCTACTCACTCAGATGAGTAGCTGCAGCACTGGCTG 384
QY 301 AGGCTGGAGTTCAGAGTGTTCGCTTGGAAAGGGCGAGTCAGAGATGACTTCTGTGTGTA 360
Db 385 AGGCTGGAGTTCAGAGTGTTCGCTTGGAAAGGGCGAGTCAGAGATGACTTCTGTGTGTA 444
QY 361 TTGACCCGCTGTGTGAACATGATGGGTGGCAGGCCAACATGATCTGTGATGATGGG 417
Db 445 TTGACCCGCTGTGTGAACATGATGGGTGGCAGGCCAACATGATCTGTGATGATGGG 501

RESULT 11
LOCUS AL554850 858 bp mRNA linear EST_16-FEB-2001
DEFINITION AL554850 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0D1087YL18 5
prime, mRNA sequence.
ACCESSION AL554850
VERSION AL554850.1 GI:12896025
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 858)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source 1..858
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1087YL18"
/clone_1lb="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 212 a 206 c 255 g 182 t 3 others
ORIGIN
Query Match 99.9%; Score 416.6; DB 9; Length 858;
Best Local Similarity 99.8%; Pred. No. 3.4e-110;
Matches 416; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAGAAGCAGCAACCACTCCAGGCGCAGCAATTTCTGTGGAAGACATCAAGC 60
Db 296 GAGAGAAGCAGCAACCACTCCAGGCGCAGCAATTTCTGTGGAAGACATCAAGC 355
QY 61 AGGAGAATTTGGACGCCGGAGATTGAGATTGCAGAGCAAGACATGCTCTGATTT 120
Db 356 AGGAGAATTTGGACGCCGGAGATTGAGATTGCAGAGCAAGACATGCTCTGATTT 415
QY 121 CACTCAGGAACAGTGTCTCAGGGGGAGAACCCCTTGCTGCTAAATAAGTGGGTGA 180
Db 416 CACTCAGGAACAGTGTCTCAGGGGGAGAACCCCTTGCTGCTAAATAAGTGGGTGA 475
QY 181 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCCGGGGCTCAGT 240
Db 476 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCCGGGGCTCAGT 535
QY 241 GCCGCTGCTCTGCTGTGTACATCTCACTCAGATGAGTAGCTGCAGCACTGGCTG 300
```

```
|||||
Db 536 GCCGCTGCTCTGCTGTGTACATCTACTCACTCAGATGAGTAGCTGCAGCACTGGCTG 595
QY 301 AGGCTGGAGTTCAGAGTGTTCGCTTGGAAAGGGCGAGTCAGAGATGACTTCTGTGTGTA 360
Db 596 AGGCTGGAGTTCAGAGTGTTCGCTTGGAAAGGGCGAGTCAGAGATGACTTCTGTGTGTA 655
QY 361 TTGACCCGCTGTGTGAACATGATGGGTGGCAGGCCAACATGATCTGTGATGATGGG 417
Db 656 TTGACCCGCTGTGTGAACATGATGGGTGGCAGGCCAACATGATCTGTGATGATGGG 712

RESULT 12
LOCUS AL527928 938 bp mRNA linear EST_13-FEB-2001
DEFINITION AL527928 LTI_NFL003_NBC3 Homo sapiens cDNA clone CS0DC027YC19 5
prime, mRNA sequence.
ACCESSION AL527928
VERSION AL527928.1 GI:12791421
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 938)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source 1..938
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DC027YC19"
/clone_1lb="LTI_NFL003_NBC3"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 241 a 208 c 275 g 212 t 2 others
ORIGIN
Query Match 99.9%; Score 416.6; DB 9; Length 938;
Best Local Similarity 99.8%; Pred. No. 3.5e-110;
Matches 416; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAGAAGCAGCAACCACTCCAGGCGCAGCAATTTCTGTGGAAGACATCAAGC 60
Db 241 GAGAGAAGCAGCAACCACTCCAGGCGCAGCAATTTCTGTGGAAGACATCAAGC 300
QY 61 AGGAGAATTTGGACGCCGGAGATTGAGATTGCAGAGCAAGACATGCTCTGATTT 120
Db 301 AGGAGAATTTGGACGCCGGAGATTGAGATTGCAGAGCAAGACATGCTCTGATTT 360
QY 121 CACTCAGGAACAGTGTCTCAGGGGGAGAACCCCTTGCTGCTAAATAAGTGGGTGA 180
Db 361 CACTCAGGAACAGTGTCTCAGGGGGAGAACCCCTTGCTGCTAAATAAGTGGGTGA 420
QY 181 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCCGGGGCTCAGT 240
Db 421 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCCGGGGCTCAGT 480
```

QY 241 GCCCGTGTCTGTGTAACATCTACTCACTCAGATGAGTAGCTGCAGCAGCTGCTG 300
|||||
Db 481 GCCCGTGTCTGTGTAACATCTACTCACTCAGATGAGTAGCTGCAGCAGCTGCTG 540
QY 301 AGCGTGAGTTGCAGTGTTCGCTTGGAGGGCGGAGTCAGAGATGACTTCTGGTGTGTA 360
|||||
Db 541 AGCGTGAGTTGCAGTGTTCGCTTGGAGGGCGGAGTCAGAGATGACTTCTGGTGTGTA 600
QY 361 TTGACCGCTGTGTGAACATGCATGGCTGGCAGGCCAACATGATCCTGATGATGGG 417
|||||
Db 601 TTGACCGCTGTGTGAACATGCATGGCTGGCAGGCCAACATGATCCTGATGATGGG 657

RESULT 13
LOCUS BG706693 701 bp mRNA linear EST 07-MAY-2001
DEFINITION 602671806F1 NIH_MGC_96 Homo sapiens CDNA clone IMAGE:4794744 5',
mRNA sequence.
ACCESSION BG706693
VERSION BG706693.1 GI:13982291
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 701)
NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shliraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10676 row: k column: 01
High quality sequence stop: 701.
Location/Qualifiers
1. 701
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4794744"
/clone_1lb="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 183 a 166 c 197 g 155 t
ORIGIN

Query Match 99.6%; Score 415.4; DB 12; Length 701;
Best Local Similarity 99.8%; Pred. No. 6.7e-110;
Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGAGAACGACGAAACCACTCCAGGGCAGCAGCAATTTCTGTGAAGAATCAAGC 60
|||||
Db 235 GAGAGAACGACGAAACCACTCCAGGGCAGCAGCAATTTCTGTGAAGAATCAAGC 294
QY 61 AGGCAAGATTGGACGGCGGAGATTGAGATTGCAGAGCAAGACATGCTCTGATT 120
|||||
Db 295 AGGCAAGATTGGACGGCGGAGATTGAGATTGCAGAGCAAGACATGCTCTGATT 354

QY 121 CACTCAGGAACGTGCTCAGGGGAGAGAGCCCTTGGCTGGTCTAAATAGTGGCTGTA 180
|||||
Db 355 CACTCAGGAACGTGCTCAGGGGAGAGAGCCCTTGGCTGGTCTAAATAGTGGCTGTA 414
QY 181 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGGGGCTGAGT 240
|||||
Db 415 CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGGGGCTGAGT 474
QY 241 GCCCGTGTCTGTGTAACATCTACTCACTCAGATGAGTAGCTGCAGCAGCTGCTG 300
|||||
Db 475 GCCCGTGTCTGTGTAACATCTACTCACTCAGATGAGTAGCTGCAGCAGCTGCTG 534
QY 301 AGCGTGAGTTGCAGTGTTCGCTTGGAGGGCGGAGTCAGAGATGACTTCTGGTGTGTA 360
|||||
Db 535 AGCGTGAGTTGCAGTGTTCGCTTGGAGGGCGGAGTCAGAGATGACTTCTGGTGTGTA 594
QY 361 TTGACCGCTGTGTGAACATGCATGGCTGGCAGGCCAACATGATCCTGATGATGGG 417
|||||
Db 595 TTGACCGCTGTGTGAACATGCATGGCTGGCAGGCCAACATGATCCTGATGATGGG 651

RESULT 14
LOCUS B1668412 857 bp mRNA linear EST 12-SEP-2001
DEFINITION 603292688F1 NIH_MGC_96 Homo sapiens CDNA clone IMAGE:5312133 5',
mRNA sequence.
ACCESSION B1668412
VERSION B1668412.1 GI:15582645
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 857)
NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shliraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11790 row: p column: 22
High quality sequence stop: 755.
Location/Qualifiers
1. 857
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5312133"
/clone_1lb="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 219 a 208 c 241 g 189 t
ORIGIN

Query Match 99.6%; Score 415.4; DB 13; Length 857;
Best Local Similarity 99.8%; Pred. No. 7.5e-110;
Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGAGACGACCAACCACTCCAGGGCAGCAGCAATTTCTGTGAGAGACATCAAGC 60
|||||
Db 242 GAGAGAACGACCAACCACTCCAGGGCAGCAGCAATTTCTGTGAGAGACATCAAGC 301
QY 61 AGGACAGATTGAGACGCCGGAGATGAGATTGACAGACAGACATGCTCTGATTT 120
|||||
Db 302 AGGACAGATTGAGACGCCGGAGATGAGATTGACAGACAGACATGCTCTGATTT 361
QY 121 CACTCAGAAACGCTGCTCAGGGGGAGAACCCCTTGGCTGCTGCTAAATAGTGGGCTGTA 180
|||||
Db 362 CACTCAGAAACGCTGCTCAGGGGGAGAACCCCTTGGCTGCTGCTAAATAGTGGGCTGTA 421
QY 181 CACACATCACAGCCAGACAGCGGTGATTGAGACACTCTGTGCCCTGGGGCTCAGT 240
|||||
Db 422 CACACATCACAGCCAGACAGCGGTGATTGAGACACTCTGTGCCCTGGGGCTCAGT 481
QY 241 GCCGCTGGTCTGCTGTAAACATCTACTCACTCAGATGAGTAGCTCAGCAGCTGGCTG 300
|||||
Db 482 GCCGCTGGTCTGCTGTAAACATCTACTCACTCAGATGAGTAGCTCAGCAGCTGGCTG 541
QY 301 AGGCTGAGTTGACAGTGTTCGCTTGGAGGGCGAGTCAGAGATGACTCTGTGGTGTGA 360
|||||
Db 542 AGGCTGAGTTGACAGTGTTCGCTTGGAGGGCGAGTCAGAGATGACTCTGTGGTGTGA 601
QY 361 TTGACCGCTGTGTGAACATGATGGGTGGCAGGCCACATGATCCTGGATGATGGGG 417
|||||
Db 602 TTGACCGCTGTGTGAACATGATGGGTGGCAGGCCACATGATCCTGGATGATGGGG 658

RESULT 15

BM449470

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

Best Local Similarity 100.0%; Pred. No. 2.1e-109;
Matches 414; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 AGAGACGACCAACCACTCCAGGGCAGCAGCAATTTCTGTGAGAGACATCAAGC 63
|||||
Db 1 AGAGACGACCAACCACTCCAGGGCAGCAGCAATTTCTGTGAGAGACATCAAGC 60
QY 64 CAGAAATTGAGACGCCGGAGATGAGATTGACAGACAGACATGCTCTGATTTGAC 123
|||||
Db 61 CAGAAATTGAGACGCCGGAGATGAGATTGACAGACAGACATGCTCTGATTTGAC 120
QY 124 TCAGAAACGCTGCTCAGGGGGAGAACCCCTTGGCTGCTGCTAAATAGTGGGCTGTAAC 183
|||||
Db 121 TCAGAAACGCTGCTCAGGGGGAGAACCCCTTGGCTGCTGCTAAATAGTGGGCTGTAAC 180
QY 184 ACATCACAGCCAGACAGCGGTGATTGAGACACTCTGTGCCCTGGGGCTCAGTGGC 243
|||||
Db 181 ACATCACAGCCAGACAGCGGTGATTGAGACACTCTGTGCCCTGGGGCTCAGTGGC 240
QY 244 GCTGCTGCTGTTGAACATCTACTCACTCAGATGAGTAGCTGACAGCACTGGCTGAG 303
|||||
Db 241 GCTGCTGCTGTTGAACATCTACTCACTCAGATGAGTAGCTGACAGCACTGGCTGAG 300
QY 304 CTGAGTTGACAGTGTTCGCTTGGAGGGCGAGTCAGAGATGACTCTGTGGTGTATG 363
|||||
Db 301 CTGAGTTGACAGTGTTCGCTTGGAGGGCGAGTCAGAGATGACTCTGTGGTGTATG 360
QY 364 ACCGCTGTGTGAACATGATGGGTGGCAGGCCACATGATCCTGGATGATGGGG 417
|||||
Db 361 ACCGCTGTGTGAACATGATGGGTGGCAGGCCACATGATCCTGGATGATGGGG 414

Search completed: April 21, 2003, 03:58:56
Job time : 568.956 secs

FEATURES

source

BM449470 1042 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT_6400881 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5493677
DEFINITION 5', mRNA sequence.
ACCESSION BM449470
VERSION BM449470.1 GI:18498510
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo;
REFERENCE 1 (bases 1 to 1042)
AUTHORS NIH-MGC http://mgi.ncl.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12117 row: m column: 06
High quality sequence stop: 694.
Location/Qualifiers
1..1042
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5493677"
/clone_11b="NIH_MGC_67"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."

BASE COUNT

267 a 219 c 309 g 240 t 7 others

ORIGIN

Query Match

99.3%; Score 414; DB 13; Length 1042;

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2003, 22:06:49 ; Search time 3345.48 Seconds

(without alignments)
11274.088 Million cell updates/sec

Title: US-09-782-051-1_COPY_549_1844

Perfect score: 1296
Sequence: 1 tccaaggcagcagcaattt.....aacctaattattacagatac 1296

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenBank:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_in:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_man:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1296	100.0	2258	9 HSU82761	U82761 Homo sapien
2	1296	100.0	2526	9 BC007576	BC007576 Homo sapi
3	1296	100.0	2552	9 BC016942	BC016942 Homo sapi
4	1296	100.0	2563	6 AX029176	AX029176 Sequence
5	1296	100.0	2583	9 BC010681	BC010681 Homo sapi
6	1296	100.0	2677	9 AF315687	AF315687 Homo sapi
7	1294.4	99.9	2510	9 HSM800298	AL049954 Homo sapi
8	1166.4	90.0	1772	10 BC018218	BC018218 Mus muscu
9	863	66.6	1987	9 AK025372	AK025372 Homo sapi
10	861.4	66.5	2030	9 BC008349	BC008349 Homo sapi
11	861.4	66.5	5025	9 AB020635	AB020635 Homo sapi
12	861.4	66.5	5052	9 BC024325	BC024325 Homo sapi
13	644	49.7	1369	9 AF035319	AF035319 Homo sapi
14	631	48.7	1792	3 AY113501	AY113501 Drosophila
15	599	46.2	1323	9 BC003631	BC003631 Homo sapi
16	486.2	37.5	40320	2 AC015387	AC015387 Drosophila
17	486.2	37.5	172372	3 AC010110	AC010110 Drosophila
18	412.6	31.8	2057	10 MUSSAHH	L32836 Mus musculu
19	411	31.7	2067	10 BC015304	BC015304 Mus muscu
20	411	31.7	190612	2 AC129093	AC129093 Mus muscu
21	406.6	31.4	52423	2 AC020395	AC020395 Drosophila
22	406.6	31.4	80423	3 DROABDB	L07835 Drosophila
23	406.6	31.4	175335	3 AC091636	AC091636 Drosophila
24	406.6	31.4	223098	3 AE003715	AE003715 Drosophila
25	406.6	31.4	338234	3 DM031961	U31961 Drosophila
26	406	31.3	1302	6 AX458351	AX458351 Sequence
27	402.4	31.0	2084	9 HUMAHCY2	M61832 Human S-ade
28	402.4	31.0	2175	9 BC010018	BC010018 Homo sapi
29	402.4	31.0	2183	9 BC011606	BC011606 Homo sapi
30	402.4	31.0	2355	9 AK097610	AK097610 Homo sapi
31	401.4	31.0	2029	10 RATAHHA	M15185 Rat S-aden
32	399	30.8	2088	5 XLA7835	AJ007835 Xenopus l
33	396.6	30.6	5050	3 DMBX200	XI3168 Drosophila
34	394.2	30.4	2129	5 XELAHH	L35559 Xenopus lae
35	375	28.9	1603	3 AY102668	AY102668 Drosophila
36	373.8	28.8	1564	3 DMAHCYGEN	X95636 D.melanogas
37	359	27.7	1617	3 AF080546	AF080546 Anopheles
38	354	27.3	1631	3 LEISADSH	M76556 L.donovan
39	346	26.7	1315	3 DDIAHHA	M19937 slime mold
40	337	26.0	384	6 AX407699	AX407699 Sequence
41	334.4	25.8	173839	2 AC118421	AC118421 Rattus no
42	322.4	24.9	37187	8 AB004537	AB004537 Schizosac
43	322.4	24.9	43757	8 SPBC8D2	AL022072 S.pombe c
44	320.8	24.8	128638	2 AC120808	AC120808 Rattus no
45	314.4	24.3	2211	6 ARI44978	ARI44978 Sequence

ALIGNMENTS

RESULT 1
HSU82761
LOCUS
DEFINITION Homo sapiens S-adenosyl homocysteine hydrolase homolog (XPYKona)
ACCESSION U82761
VERSION U82761
KEYWORDS mRNA, complete cds.
SOURCE
ORGANISM Homo sapiens.
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 2258)
Cleever, J.E., Afzal, V., Feeney, L., McDowell, M., Sadinskl, W.,
Volpe, J.P.G., Busch, D.B., Coleman, D.M., Ziffer, D.W., Yu, Y.,

Pred. No. 1s the number of results predicted by chance to have a

TITLE Nagasawa,H. and Little,J.B.
Increased ultraviolet sensitivity and chromosomal instability
related to P53 function in the xeroderma pigmentosum variant
JOURNAL Cancer Res. 59 (5), 1102-1108 (1999)
MEDLINE 99168517
PUBMED 10070969
REFERENCE 2 (bases 1 to 2258)
AUTHORS Volpe,J.P.G., McDowell,M., Jostes,R.F., Afzal,V., Sadinski,W.,
Trask,B.J., Legerski,R. and Cleaver,J.E.
TITLE Complementatation of chromosomal instability in the xeroderma
pigmentosum variant by a gene on human chromosome 1 with homology
to S-adenosyl homocysteine hydrolase
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 2258)
AUTHORS Volpe,J.P.G., McDowell,M. and Cleaver,J.E.
TITLE Direct Submission
JOURNAL Submitted (19-DEC-1996) Dermatology, UCSF, 3rd and Parnassus, Box
0750, San Francisco, CA 94143, USA
FEATURES
SOURCE location/Qualifiers
1. 2258
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/cell_type="fibroblasts"
/tissue_type="skin"
1. 2258
/gene="XPVKona"
48. 1550
/note="Xeroderma pigmentosum variant"
/codon_start=1
/product="S-adenosyl homocysteine hydrolase homolog"
/protein_id="AAC01960.1"
/db_xref="GI:2852125"
/translation="MATVTAKPKQIQFADDMQEFTEPTKGRSLRSISQSTDS
YSSAASYTDSDEVSPEKQQTNSKSSNFCVKNIKQAEFGREIEIAQDMALIS
LKRAGEKPLAGAKIVGCTHITAGTAVLIETICALGACRMSACNISTONEVAAL
AEAGVAFAMKGESEDDFWWCIDRCVNDGQANMILDDGDLTHVYKKYPNFKKI
RGIVEESTGVHRLYQLSKAGKLCVPAMVNDVTKQKFDNLCCRESILDLKRTTD
VMFGKQYVVCYGEYGVKGCACALKAIAIVYTEIDPICALQACMDGFRVVKLNEVI
RQVDVITCTGNKNVTRHLDPMKNSCIVCMGHSNTEIDVTSLEPDLTWERVRSO
VDHVTWDEGKRVLLAEGRLNLSCTVPTFVLSITATTOALALIELYNAPEGRYKOD
VYLLPKMDEYVASLHLPSEDAHLTELTDQAKYLGINKNGPRPNYYRY"
BASE COUNT 603 a 513 c 542 g 600 t
ORIGIN
Query Match 100.0%; Score 1296; DB 9; Length 2258;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCCAAGGGCAGCAGCAATTTCTGTGTAAGAACAATCAAGCAGGAGAAATTTGACGCCGG 60
DB 252 TCCAAGGGCAGCAGCAATTTCTGTGTAAGAACAATCAAGCAGGAGAAATTTGACGCCGG 311
QY 61 GAGATTGAGATTGAGAGCAAGACATGTCTGCTGATTCTCACTCAGGAACGCTGCTCAG 120
DB 312 GAGATTGAGATTGAGAGCAAGACATGTCTGCTGATTCTCACTCAGGAACGCTGCTCAG 371
QY 121 GGGGAGAGCCCTGGCTGCTCTAAATAGTGGGCTGTACACACATCACAGCCAGACA 180
DB 372 GGGGAGAGCCCTGGCTGCTCTAAATAGTGGGCTGTACACACATCACAGCCAGACA 431
QY 181 GCGGTGTGATTGAGACACTCTGTCCCTGGGGCTCAGTCCGCTGCTGCTGTAAAC 240
DB 432 GCGGTGTGATTGAGACACTCTGTCCCTGGGGCTCAGTCCGCTGCTGCTGTAAAC 491
QY 241 ATCTACTCACTCAGATGAAGTAGCTGACGACACTGGCTGAGGCTGAGGTGCAAGTTC 300
DB 492 ATCTACTCACTCAGATGAAGTAGCTGACGACACTGGCTGAGGCTGAGGTGCAAGTTC 551
QY 301 GCTTGGAAGGGCGAGTCAGAGATGACTTCTGCTGCTGCTATTTGACCGCTGTGTGAACATG 360
DB 552 GCTTGGAAGGGCGAGTCAGAGATGACTTCTGCTGCTGCTATTTGACCGCTGTGTGAACATG 611

QY 361 GATGGGTGCGAGGCCAATGATCTCTGATGATGGGGAGACTTAACCCACTGGGTTAT 420
DB 612 GATGGGTGCGAGGCCAATGATCTCTGATGATGGGGAGACTTAACCCACTGGGTTAT 671
QY 421 AAGAAGTATCCAAACGCTTTAAGAAGATCCGAGGCAATTGTGAAGAGAGCGTGAAGT 480
DB 672 AAGAAGTATCCAAACGCTTTAAGAAGATCCGAGGCAATTGTGAAGAGAGCGTGAAGT 731
QY 481 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTCTCCGGCCATGAAGCTC 540
DB 732 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTCTCCGGCCATGAAGCTC 791
QY 541 AATGATCTCTTACCACAGAGAGTTGATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
DB 792 AATGATCTCTTACCACAGAGAGTTGATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTG 851
QY 601 GATGGCTGAAGAGGACACAGATGATGATGTTGTTGGGAACAAGTGTGTGTGTGTC 660
DB 852 GATGGCTGAAGAGGACACAGATGATGATGTTGTTGGGAACAAGTGTGTGTGTGTC 911
QY 661 TATGCTGAGTAGGCAAGGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
DB 912 TATGCTGAGTAGGCAAGGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 971
QY 721 ATTACGGAATCGACCCCACTGTGCTCTGACAGGCTGCAATGATGGTTCAGGGGTGTA 780
DB 972 ATTACGGAATCGACCCCACTGTGCTCTGACAGGCTGCAATGATGGTTCAGGGGTGTA 1031
QY 781 AAGCTAATGAAGTCACTCCGCAAGTGCATGCTGCTAATTAATCTGCACAGGAATAAGAT 840
DB 1032 AAGCTAATGAAGTCACTCCGCAAGTGCATGCTGCTAATTAATCTGCACAGGAATAAGAT 1091
QY 841 GTAGTGACACGGGACCTGCTGATGCAAAAAACAGTGTATCGTATGCAATATGGGC 900
DB 1092 GTAGTGACACGGGACCTGCTGATGCAAAAAACAGTGTATCGTATGCAATATGGGC 1151
QY 901 CACTCCAAACAGAAATCGATGTGACACAGCCTCCGCACTCCGCACTGACGTGGAGCGA 960
DB 1152 CACTCCAAACAGAAATCGATGTGACACAGCCTCCGCACTCCGCACTGACGTGGAGCGA 1211
QY 961 GTAGCTTCTCAGGTGACATGTCATCTGCGCAGATGGCAACGAGTGTCTCTCTGCGCA 1020
DB 1212 GTAGCTTCTCAGGTGACATGTCATCTGCGCAGATGGCAACGAGTGTCTCTCTGCGCA 1271
QY 1021 GAGGCTGCTCTACTCAATTTGAGCTGCTCCACAGTTCACACCTTTGTTCTGCTCATACA 1080
DB 1272 GAGGCTGCTCTACTCAATTTGAGCTGCTCCACAGTTCACACCTTTGTTCTGCTCATACA 1331
QY 1081 GCCACAACACAGGCTTTGGCACTGATAGACTCTATAATGCAACCCGAGGGCGATACAG 1140
DB 1332 GCCACAACACAGGCTTTGGCACTGATAGACTCTATAATGCAACCCGAGGGCGATACAG 1391
QY 1141 CAGGATGTGACTTGTCTTAAGAAATGATGAATAGTTCAGCTTGCCATCTGCCA 1200
DB 1392 CAGGATGTGACTTGTCTTAAGAAATGATGAATAGTTCAGCTTGCCATCTGCCA 1451
QY 1201 TCATTGTATGCCCACTTACAGAGCTGACAGATGACCAACCAAAATATCTGGGACTCAAC 1260
DB 1452 TCATTGTATGCCCACTTACAGAGCTGACAGATGACCAACCAAAATATCTGGGACTCAAC 1511
QY 1261 AAAAATGGCCATTCAACCTAATTATTACAGATAC 1296
DB 1512 AAAAATGGCCATTCAACCTAATTATTACAGATAC 1547
RESULT 2.
LOCUS BC007576 2526 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, S-adenosylhomocysteine hydrolase-like 1, clone
ACCESSION BC007576
VERSION BC007576.1 GI:14043176

KEYWORDS: MGC.
SOURCE: Homo sapiens.
ORGANISM: Homo sapiens
REFERENCE: 1 (bases 1 to 2526)
AUTHORS: Strausberg, R.
TITLE: Direct Submission
JOURNAL: Submitted (10-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK: NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT: Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubln Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guln, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Lisa Prabh, Parvaneh Saeedi, Jacqueline Scheln, Duane Smalus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAL Plate: 22 Row: h Column: 21
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.

FEATURES

Source
1. 2526
/organism="Homo sapiens"
/db_xref="locusid:10768"
/db_xref="taxon:9606"
/clone="MGC:15558 IMAGE:3139729"
/tissue_type="Placenta, choriocarcinoma"
/clone_id="NIH_MGC_21"
/lab_host="DH10B-R"
/note="Vector: POTB7"
308. 1810
/codon_start=1
/product="S-adenosylhomocysteine hydrolase-like 1"
/protein_id="AAH07576.1"
/db_xref="GI:14043177"
/translation="MATVTKAPKKQIQFADMQEFTKPTKGRSLRSISOSTDS
YSSAASYTDSDEVSPEKQOTNSGSSNFCVKNIKQAEFGREIEIAEDMSALIS
LRKRAQGERPLAGAKIVGCTHTAIVLIEIICALAGCWSACNISTQNEVAAL
AEAGVAVFAMGSESDDEFWCIDRCVNMDCWQANMILDDGDLTHWVYKKYNEVKKI
RGIVEESVTGVHRLIQLSKAGKLCVPAMVNDSTKOKFDNLCCRESILDLKRTD
VMFGKOVVCGYGEVKGCCALKALGAIIVYIEIDPICALQACMDGFRVYKLINEVI
ROVDVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVRSO
VDHVIMPDGRVLLAEGRLINLSCSTVPTFVLSITATTOALALIELYNAPGCRYKQD
VYLLPKRMDEYVASHLPSFDAHLTELTDQAKYILGLNKNGPKNRYRY"

BASE COUNT 644 a 592 c 660 g 630 t
ORIGIN

Query Match 100.0%; Score 1296; DB 9; Length 2526;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCAAGGCGACGACATTTCTGTGAGAGAACATCAAGCAGGCGAATTGGAGCGCCG 60
Db 512 TCCAAGGCGACGACATTTCTGTGAGAGAACATCAAGCAGGCGAATTGGAGCGCCG 571
QY 61 GAGATTGAGATTGACAGACAGACATGCTGCTGTGATTTCACTCAGGAACGCTGCTAG 120

Db 572 GAGATTGAGATTGACAGACAGACATGCTGCTGTGATTTCACTCAGGAACGCTGCTAG 631
QY 121 GGGAGAGAGCCCTTGGCTGTGCTAAATAGTGGGCTGTACACATCACAGCCACAGA 180
Db 632 GGGAGAGAGCCCTTGGCTGTGCTAAATAGTGGGCTGTACACATCACAGCCACAGA 691
QY 181 GCGGTGTGATTGAGACACTCTGTGCTGCGGCTGCGGCTCAGTCCGCTGTCTGTGTAAC 240
Db 692 GCGGTGTGATTGAGACACTCTGTGCTGCGGCTGCGGCTCAGTCCGCTGTCTGTGTAAC 751
QY 241 ATCTACTCAACTCAGAAATGAAATGAGTGTGACAGCACTGCTGAGCTGAGTTGACAGTTTC 300
Db 752 ATCTACTCAACTCAGAAATGAAATGAGTGTGACAGCACTGCTGAGCTGAGTTGACAGTTTC 811
QY 301 GCTTGAAGGCGGAGTCAAGAGATGACTTCTGTGTGTATGACCCGCTGTGTGAACATG 360
Db 812 GCTTGAAGGCGGAGTCAAGAGATGACTTCTGTGTGTATGACCCGCTGTGTGAACATG 871
QY 361 GATGGGTGGCAGGCGCAACATGATCTGTGATGATGGGGAGACTTAACCCACTGGGTTTAT 420
Db 872 GATGGGTGGCAGGCGCAACATGATCTGTGATGATGGGGAGACTTAACCCACTGGGTTTAT 931
QY 421 AAGAAGTATCCAAAGCTGTTTAAGAGATCCGAGGCTATGTTGAGAGAGCGTACTGCT 480
Db 932 AAGAAGTATCCAAAGCTGTTTAAGAGATCCGAGGCTATGTTGAGAGAGCGTACTGCT 991
QY 481 GTTACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTCTCCGCGCATGAAGCTC 540
Db 992 GTTACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTCTCCGCGCATGAAGCTC 1051
QY 541 AATGATCTCTTACCAACAGAGATTGATTAATCTGTACTGCTGCGGAGAAATCCATTG 600
Db 1052 AATGATCTCTTACCAACAGAGATTGATTAATCTGTACTGCTGCGGAGAAATCCATTG 1111
QY 601 GATGGCTGAAGAGAGACACAGATGTGATGTTGTTGGGAAACAACTGTGTGTGCTGC 660
Db 1112 GATGGCTGAAGAGAGACACAGATGTGATGTTGTTGGGAAACAACTGTGTGTGCTGC 1171
QY 661 TATGTTGAGTAGGCAAGGCTGTGCTGTCTCTCAAGCTCTGTGAGCAATGTCTTAC 720
Db 1172 TATGTTGAGTAGGCAAGGCTGTGCTGTCTCTCTCAAGCTCTGTGAGCAATGTCTTAC 1231
QY 721 ATTACCGAATCGACCCCATCTGTGCTGTGCAAGGCTGCATGAGTGGGTTGAGGTTGTA 780
Db 1232 ATTACCGAATCGACCCCATCTGTGCTGTGCAAGGCTGCATGAGTGGGTTGAGGTTGTA 1291
QY 781 AAGCTAATGAAGTCATCCGCGCAAGTCATGCTGTAATACTGACACAGGAATAAGAAAT 840
Db 1292 AAGCTAATGAAGTCATCCGCGCAAGTCATGCTGTAATACTGACACAGGAATAAGAAAT 1351
QY 841 GTAGTGACAGGGAGACACTTGGATGCGATGAATAAAGAGTTGATGATGCAATATGGGC 900
Db 1352 GTAGTGACAGGGAGACACTTGGATGCGATGAATAAAGAGTTGATGATGCAATATGGGC 1411
QY 901 CACTCCAAACAGAAATGATGTGACAGCCTCCGCACTCCGAGCTGACGTGGAGCGA 960
Db 1412 CACTCCAAACAGAAATGATGTGACAGCCTCCGCACTCCGAGCTGACGTGGAGCGA 1471
QY 961 GTACGTTCTCAGGTGAGCAGCATGTGCTGCGCAGATGGCAAAAGAGTTGCTCTCTGGA 1020
Db 1472 GTACGTTCTCAGGTGAGCAGCATGTGCTGCGCAGATGGCAAAAGAGTTGCTCTCTGGA 1531
QY 1021 GAGGCTGCTACTCAATTTGAGCTGCTCCACAGTTCCACCTTGTCTGTCCATCACA 1080
Db 1532 GAGGCTGCTACTCAATTTGAGCTGCTCCACAGTTCCACCTTGTCTGTCCATCACA 1591
QY 1081 GCCACAACACAGGCTTTGGCACTGATGAACCTTATATGACACCCGAGGGCGATAGAG 1140
Db 1592 GCCACAACACAGGCTTTGGCACTGATGAACCTTATATGACACCCGAGGGCGATAGAG 1651
QY 1141 CAGGATGTGACTTGTCTCTTAAGAAATGATGAATACGTTGCGAGCTTGATCTGCCA 1200
Db 1652 CAGGATGTGACTTGTCTCTTAAGAAATGATGAATACGTTGCGAGCTTGATCTGCCA 1711

Oy 1201 TCATTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGACTCAAC 1260
|||||
Db 1712 TCATTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGACTCAAC 1771
Oy 1261 AAAAAAGGCCATTCACACCTAATTAATACAGATAC 1296
|||||
Db 1772 AAAAAAGGCCATTCACACCTAATTAATACAGATAC 1807

RESULT 3

BC016942 2552 bp mRNA linear PRI-09-NOV-2001

LOCUS Homo sapiens, S-adenosylhomocysteine hydrolase-like 1, clone

DEFINITION MGC:21453 IMAGE:3450568, mRNA, complete cds.

ACCESSION BC016942

VERSION BC016942.1 GI:16877386

KEYWORDS MGC.

SOURCE Homo sapiens.

ORGANISM Homo sapiens.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 2552)

AUTHORS Strausberg, R.

JOURNAL Direct Submission

Submitted (05-NOV-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>

Contact: MGC help desk

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome

Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www.shgc.stanford.edu>

Contact: (Dickson, Mark) medepaxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,

R. M.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>

Series: IRAK Plate: 20 Row: a Column: 15

This clone was selected for full length sequencing because it

passed the following selection criteria: Similarity but not

identity to protein.

FEATURES

source location/Qualifiers

1. 2552

/organism="Homo sapiens"

/db_xref="locusid:10768"

/db_xref="taxon:9606"

/clone="MGC:21453 IMAGE:3450568"

/tissue_type="Placenta, choriochorionoma"

/clone_lib="NH_MGC_10"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

338. 1840

/codon_start=1

/product="S-adenosylhomocysteine hydrolase-like 1"

/protein_id="AAH16942.1"

/db_xref="GI:16877387"

/translation="MATVTAPKPKKIOFADDMOEFKPTKGRSLRSISOSTDS

YSSASYSDDSDDEVPREKQOTNSKSSNECVKNIKQAEFGREIEIAEDMSALIS

LKRAOGEKPLAGKIVGCTHITATVLIETICAIAGACRSACNISTONEVAAL

AEAGVAFAMKGESEDDFWCIDRCVNMGMQANMLDGGDLTHWYKKIPNYEKKI

RGIVESVTGVHRLYOLSKAGKICVPAMVNDSTYKOKEDNLCCRESILDKRTD

VMFGKQVVCVCGYGVKGCACALALGAIYITIEDPICALQACMDGFRVVKINEVI

RQDVVITCTGNKNVYREHLDMKNSCIVCNMGHSTIEDVTSRTPELFWERVRSD

VDHVIWPDGKRVLLAEGRLLNLSCSTVPTFVLSITATQALALIELYNAPEGRYKQD

VLLPKMDEVASLHPSFDAHLELTDQAKYLGINKNGPKPNYYRY"

BASE COUNT 644 a 600 c 671 g 637 t

ORIGIN

Query Match 100.0%; Score 1296; DB 9; Length 2552;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TCCAAAGGCGACGACCAATTTCTGTGTGAAGAACATCAAGCAGGCGAGAAATTTGGACGCCG 60
|||||
Db 542 TCCAAAGGCGACGACCAATTTCTGTGTGAAGAACATCAAGCAGGCGAGAAATTTGGACGCCG 601

Oy 61 GAGATTGAGATTGACAGACGACATGCTGCTGATTTCACTCAGGAAAGTGTCTAG 120
|||||

Db 602 GAGATTGAGATTGACAGACGACATGCTGCTGATTTCACTCAGGAAAGTGTCTAG 661

Oy 121 GGGAGAGAGCCCTTGCTGCTGCTAAATAGTGGGCTGTACACACATCACAGCCAGACA 180
|||||

Db 662 GGGAGAGAGCCCTTGCTGCTGCTAAATAGTGGGCTGTACACACATCACAGCCAGACA 721

Oy 181 GCGGTGTGATTGAGACACTCTGTCCTGCGGCGCTCAGTGCCTGCTGTGTAAC 240
|||||

Db 722 GCGGTGTGATTGAGACACTCTGTCCTGCGGCGCTCAGTGCCTGCTGTGTAAC 781

Oy 241 ATCTACTCAACTCAGAAATGAGTAGCTGACGACTGGCTGAGGCTGAGTGTGAGTGTG 300
|||||

Db 782 ATCTACTCAACTCAGAAATGAGTAGCTGACGACTGGCTGAGGCTGAGTGTGAGTGTG 841

Oy 301 GCTTGAAGGCGAGTCAAGAGATGACTTGTGTGTGTATGACCGCTGTGTGAACATG 360
|||||

Db 842 GCTTGAAGGCGAGTCAAGAGATGACTTGTGTGTGTATGACCGCTGTGTGAACATG 901

Oy 361 GATGGGTGCGACCAACATGATCTGTGATGATGGGGAGAGACTTAACCACTGGGTTAT 420
|||||

Db 902 GATGGGTGCGACCAACATGATCTGTGATGATGGGGAGAGACTTAACCACTGGGTTAT 961

Oy 421 AAGAAGTATCCAAACGTGTTAAGAAGATCCAGGCAATGTGAAGAGAGCGTGACTGT 480
|||||

Db 962 AAGAAGTATCCAAACGTGTTAAGAAGATCCAGGCAATGTGAAGAGAGCGTGACTGT 1021

Oy 481 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTGCCGCAATGAACGTC 540
|||||

Db 1022 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTGCCGCAATGAACGTC 1081

Oy 541 AATGATTCCTGTACCAACAGAGATTGATTAATCTGTACTGCTGCGGCAATTCATTG 600
|||||

Db 1082 AATGATTCCTGTACCAACAGAGATTGATTAATCTGTACTGCTGCGGCAATTCATTG 1141

Oy 601 GATGGCTGAAGAGGACACACAGATGTGATTTGGTGGGAACAAGTGTGTGTGTCG 660
|||||

Db 1142 GATGGCTGAAGAGGACACACAGATGTGATTTGGTGGGAACAAGTGTGTGTGTCG 1201

Oy 661 TATGTTGAGTAGGCAAGGGCTGTGTGCTCTCAAAAGCTCTTGAGCAATTTGCTAC 720
|||||

Db 1202 TATGTTGAGTAGGCAAGGGCTGTGTGCTCTCAAAAGCTCTTGAGCAATTTGCTAC 1261

Oy 721 ATTACCGAATCGACCCCATCTGTGCTGTGAGGCTGATGATGGGTGAGGTTGTA 780
|||||

Db 1262 ATTACCGAATCGACCCCATCTGTGCTGTGAGGCTGATGATGGGTGAGGTTGTA 1321

Oy 781 AAGCTAAATGAGTCAATCCGCAAGTGTGATGCTAATACTTGACAGAAATAAGAT 840
|||||

Db 1322 AAGCTAAATGAGTCAATCCGCAAGTGTGATGCTAATACTTGACAGAAATAAGAT 1381

Oy 841 GTAGTGACAGGAGACACTGTGATCGCATGAATAACAGTGTATCGTATGCAATATGGGC 900
|||||

Db 1382 GTAGTGACAGGAGACACTGTGATCGCATGAATAACAGTGTATCGTATGCAATATGGGC 1441

Oy 901 CACTCCAACACAGAAATCGATGTGACCAAGCTCCGCACTCCGAGCTGACGTGGAGCGA 960
|||||

Db 1442 CACTCCAACACAGAAATCGATGTGACCAAGCTCCGCACTCCGAGCTGACGTGGAGCGA 1501

Oy 961 GTACGTTCTCAGGTGACCATGTCACTGCGCAGATGGCAACGAGTGTCTCTGCA 1020
|||||

Db 1502 GTACGTTCTCAGGTGACCATGTCACTGCGCAGATGGCAACGAGTGTCTCTGCA 1561

QY 1021 GAGGTCGTCTACTCAATTTGAGCTGCTCCACAGTTCACCTTTGTCTGTCCATCACA 1080
|||||
Db 1562 GAGGTCGTCTACTCAATTTGAGCTGCTCCACAGTTCACCTTTGTCTGTCCATCACA 1621
|||||
QY 1081 GCCACACACAGAGCTTTGGCAGCTGATAGAACTCTATATGCAACCCGAGGGCGATACAAAG 1140
|||||
Db 1622 GCCACACACAGAGCTTTGGCAGCTGATAGAACTCTATATGCAACCCGAGGGCGATACAAAG 1681
|||||
QY 1141 CAGCATGTGTACTGCTCTCTTAAGAAAATGATGAATACGTTGCCAGCTTGACATCTGCCA 1200
|||||
Db 1682 CAGCATGTGTACTGCTCTCTTAAGAAAATGATGAATACGTTGCCAGCTTGACATCTGCCA 1741
|||||
QY 1201 TCATTTGATGCCCCACCTTACAGAGCTGACAGATGACCAAGAAAATATCTGGGACTCAAC 1260
|||||
Db 1742 TCATTTGATGCCCCACCTTACAGAGCTGACAGATGACCAAGAAAATATCTGGGACTCAAC 1801
|||||
QY 1261 AAAAATGGCCATTCAAAACCTAATTATTACAGATAC 1296
|||||
Db 1802 AAAAATGGCCATTCAAAACCTAATTATTACAGATAC 1837
|||||

RESULT 4
AX029176 2563 bp DNA linear PAT 16-SEP-2000
LOCUS AX029176
DEFINITION Sequence 1 from Patent WO9814562.
ACCESSION AX029176
VERSION AX029176.1 GI:10190060
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 2563)
AUTHORS Hart,D.N.
TITLE Enzyme having s-adenosyl-l-homocysteine hydrolase (ahcy) type
activity
JOURNAL Patent: WO 9814562-A.1 09-APR-1998;
HART,DEREK NIGEL JOHN (NZ);
FEATURES
source location/Qualifiers
1. 2563
/organism="Homo sapiens"
/db_xref="taxon:9606"
CDS
<3..1847
/note="Open reading frame extends without a stop codon for
the full 5' nucleotide sequence. The initiation codon has
yet to be identified."
/codon_start=1
/protein_id="CAC09285.1"
/db_xref="GI:10190061"
/translation="RGQVGARSCTFWSGHRCPALGCRDTDKAMATAPQKPTQLDA
GAGRVGDRVSEGAARAGRAPERGGGSAAGAGGSMPLDPLGVEELKO
AKEIEDAEKSEMATVTKAPKKQIQFADDMQETKPTKGRSLRSISGSDTSYS
SAASYTSSDDEVSPEKQOTNSQSNFCVKNKIOAEFGRREIEIAEODMSALISLR
KRAOGEKPLAGAKIVGCTHTTAQTAVLIETLALGAQCRWSACNISTONEVAALAE
AGVAFAWKESEDDFWCIDRCVNMGMNMLDDGGDLTHVYKKYPPNVEKKIRG
IVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLCCRESILDGKRTDVM
FGKQVVVCGYGEVKGCCCAALKALGAIVYITIEDICALQACMDGFRVYKINEVIRQ
VDVITCTGKNVYTRHEIDRMKNSCIYCNMGSNTEIDVTSLRTPELTWERVRSQVD
HVIMPDGKRVLLAEGRLNLSCSTVPITVLSITATQALALIELYNAPEGRRYKQDVI
LLPKMDEYVASLHLSFSDAHLTELTDQAKYLGINKGPKPNYRY"

BASE COUNT 646 a 604 c 677 g 636 t
ORIGIN

Query Match 100.0%; Score 1296; DB 6; Length 2563;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCAAAGGGCAGCAGCAATTTCTGTGTGAAGAACATCAAGCAGCAGAAATTTGGAGCGCCG 60
|||||
Db 549 TCCAAAGGGCAGCAGCAATTTCTGTGTGAAGAACATCAAGCAGCAGAAATTTGGAGCGCCG 608
|||||
QY 61 GAGATGTGATGTGAGAGCAAGACATGTCTGCTGTGATTTCACTCAGGAAACGTGCTCAG 120
|||||

Db 609 GAGATGTGATGTGAGAGCAAGACATGTCTGCTGTGATTTCACTCAGGAAACGTGCTCAG 668
|||||
QY 121 GGGAGAGCCCTTGCGTGTGCTAAATAAGTGGGCTGTACACACATCAGCCAGACA 180
|||||
Db 669 GGGAGAGCCCTTGCGTGTGCTAAATAAGTGGGCTGTACACACATCAGCCAGACA 728
|||||
QY 181 GCGGTGTGATTTGAGACACTCTGTGCCCCGCTGAGGGGCTCAGTGGCCGCTGCTGTGTAAC 240
|||||
Db 729 GCGGTGTGATTTGAGACACTCTGTGCCCCGCTGAGGGGCTCAGTGGCCGCTGCTGTGTAAC 788
|||||
QY 241 ATCTACTCACTCAGAAATGAAGTAGCTGACAGCACTGGCTGAGGCTGAGGTTGCACTGTTTC 300
|||||
Db 789 ATCTACTCACTCAGAAATGAAGTAGCTGACAGCACTGGCTGAGGCTGAGGTTGCACTGTTTC 848
|||||
QY 301 GCTTGAAGGCGAGTCAAGAGATGACTTCTGTGTGTATTGACCGCTGTGTGAACATG 360
|||||
Db 849 GCTTGAAGGCGAGTCAAGAGATGACTTCTGTGTGTATTGACCGCTGTGTGAACATG 908
|||||
QY 361 GATGGTGGCAGGCCAACATGATCTCGATGATGGGGAGAGACTTAACCCACTGGGTTAT 420
|||||
Db 909 GATGGTGGCAGGCCAACATGATCTCGATGATGGGGAGAGACTTAACCCACTGGGTTAT 968
|||||
QY 421 AAGAAGTATCCAAACGTGTTAAGAAGATCCGAGGCACTGTGGAAGAGACCGTGACTGCT 480
|||||
Db 969 AAGAAGTATCCAAACGTGTTAAGAAGATCCGAGGCACTGTGGAAGAGACCGTGACTGCT 1028
|||||
QY 481 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTGTCCGGCCATGAACGTC 540
|||||
Db 1029 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTGTCCGGCCATGAACGTC 1088
|||||
QY 541 AATGATTTCTTACCAACAGAGTTTGAATTACTGTACTGCTGCCAGAAATCCATTTTG 600
|||||
Db 1089 AATGATTTCTTACCAACAGAGTTTGAATTACTGTACTGCTGCCAGAAATCCATTTTG 1148
|||||
QY 601 GATGGCTGAAGAGACACACAGATGTGATGTTGGTGGAAACAAGTGGTGTGTGGC 660
|||||
Db 1149 GATGGCTGAAGAGACACACAGATGTGATGTTGGTGGAAACAAGTGGTGTGTGGC 1208
|||||
QY 661 TATGCTGAGGTAGGCAAGGGCTGCTGTGCTCTCAAGCTCTTGGAGCAATGTCTAC 720
|||||
Db 1209 TATGCTGAGGTAGGCAAGGGCTGCTGTGCTCTCAAGCTCTTGGAGCAATGTCTAC 1268
|||||
QY 721 ATTACCGAAATCGAACCCCATCTGTGCTGTGCAAGGCTGCATGATGGGTTGAGGGTGA 780
|||||
Db 1269 ATTACCGAAATCGAACCCCATCTGTGCTGTGCAAGGCTGCATGATGGGTTGAGGGTGA 1328
|||||
QY 781 AAGCTAAATGAAGTCAATCCGGCAAGTGTGCTAATAACTTGCACAGGAATAAGAAAT 840
|||||
Db 1329 AAGCTAAATGAAGTCAATCCGGCAAGTGTGCTAATAACTTGCACAGGAATAAGAAAT 1388
|||||
QY 841 GTAGTACAGGGAGCACTTGGATGCGATGAAGAAAACAGTTGTATCGTATGCAATATGGGC 900
|||||
Db 1389 GTAGTACAGGGAGCACTTGGATGCGATGAAGAAAACAGTTGTATCGTATGCAATATGGGC 1448
|||||
QY 901 CACTCCACACAGAAATCGATGTGACCAAGCTCCGCACTCCGGAGCTGACGTGGAGCGA 960
|||||
Db 1449 CACTCCACACAGAAATCGATGTGACCAAGCTCCGCACTCCGGAGCTGACGTGGAGCGA 1508
|||||
QY 961 GTACGTTCTCAGGTGAGCAGCATGTCTGCGCAGATGGCAACAGAGTGTCTCTCTGCGCA 1020
|||||
Db 1509 GTACGTTCTCAGGTGAGCAGCATGTCTGCGCAGATGGCAACAGAGTGTCTCTCTGCGCA 1568
|||||
QY 1021 GAGGTCGTCTACTCAATTTGAGCTGCTCCACAGTTCACCTTTGTCTGTCCATCACA 1080
|||||
Db 1569 GAGGTCGTCTACTCAATTTGAGCTGCTCCACAGTTCACCTTTGTCTGTCCATCACA 1628
|||||
QY 1081 GCCACACACAGAGCTTTGGCAGCTGATAGAACTCTATATGCAACCCGAGGGCGATACAAAG 1140
|||||
Db 1629 GCCACACACAGAGCTTTGGCAGCTGATAGAACTCTATATGCAACCCGAGGGCGATACAAAG 1688
|||||
QY 1141 CAGATGTGATGTGAGAGCAAGACATGTCTGCTTAAGAAAATGATGAATACGTTGCCAGCTTGCACTGCCA 1200
|||||

Db 1689 CAGGATGTACTTCTTCTTAAGAAATGATGATACGTGGCCAGCTTGCACTGCCA 1748

QY 1201 TCATTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGACTCAAC 1260

Db 1749 TCATTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGACTCAAC 1808

QY 1261 AAAATGGGCATCAACCTAATTATATACAGATAC 1296

Db 1809 AAAATGGGCATCAACCTAATTATATACAGATAC 1844

RESULT 5

BC010681 2583 bp mRNA linear PRI 12-JUL-2001

LOCUS Homo sapiens, S-adenosylhomocysteine hydrolase-like 1, clone

DEFINITION MGC:8936 IMAGE:3853747, mRNA, complete cds.

ACCESSION BC010681

VERSION BC010681.1 GI:14715037

KEYWORDS MGC.

SOURCE Homo sapiens.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS 1 (bases 1 to 2583)

TITLE Strausberg, R.

JOURNAL Direct Submission

Submitted (10-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: cgabs-r@mail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: villalobcm.tmc.edu

Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>

Series: IRAC Plate: 13 Row: b Column: 21

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 5729723.

FEATURES

source

1. .2583

/organism="Homo sapiens"

/db_xref="LocusID:10768"

/db_xref="taxon:9606"

/clone="MGC:8936 IMAGE:3853747"

/tissue_type="Colon, adenocarcinoma"

/clone_lib="NIH_MGC_65"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

367. .1869

/codon_start=1

/product="S-adenosylhomocysteine hydrolase-like 1"

/protein_id="AAH10681.1"

/db_xref="GI:14715038"

/translation="MATVTKAPKQIQFADDMQEFKPTKTRKSLRSISQSTDS YSSAASYTSDSEVSPREKQOTNSKSSNECVKNIKOAFGRREIEIAEODMSALIS LKRAAGEKPLAGAKIVGCTHTTAQAVLIEFLCALGACRWSACNITSTONEVAAL AEAGVAFAWKGESEDEDFWMCIDRCVNMNDGQANMIDGSDLTHWVKKYPNEFKI RGIYESVTGVHRLYQLSKAGKLCVPAMNNDVTKQKFDNLCCRESILDGKRTED VMFGKQVAVCGYGEVGKCCALALGAIIVYTEIDPICALQACMDGFRVVKLNEVI ROYDVITCTGNKNVTRHEDRMKNSCIVCMGHSNTEIDVTSRTPELTWERVRSO VDHVIMPDGKRIVLLAEGRLNLSGSTVPTFVLSTATQALALIELYNAPGRIKOD

CDS

BASE COUNT 648 a 606 c 688 g 641 t

ORIGIN

Query Match 100.0%; Score 1296; DB 9; Length 2583;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCAAGGCGACGACCAATTTCTGTGTGAAGAACATCAAGCAGCAGAAATTTGGACGCCGG 60

Db 571 TCCAAGGCGACGACCAATTTCTGTGTGAAGAACATCAAGCAGCAGAAATTTGGACGCCGG 630

QY 61 GAGATTGAGATTGACAGCAAGACATGTCTGCTCTGATTTCACTCAGAAACGTGCTAG 120

Db 631 GAGATTGAGATTGACAGCAAGACATGTCTGCTCTGATTTCACTCAGAAACGTGCTAG 690

QY 121 GGGAGAAAGCCCTTGGCTGTGCTAAATATAGTGGCTGTACACACATACAGCCAGACA 180

Db 691 GGGAGAAAGCCCTTGGCTGTGCTAAATATAGTGGCTGTACACACATACAGCCAGACA 750

QY 181 GCGGTGTGATTGAGACACTGTGTGCCCTGGGGGCTCAGTGCCGCTGCTGTGTAAC 240

Db 751 GCGGTGTGATTGAGACACTGTGTGCCCTGGGGGCTCAGTGCCGCTGCTGTGTAAC 810

QY 241 ATCTACTCAACTCAGAAATGAAGTAGCTGACAGCACTGGCTGAGGCTGAGTGCAGTGTTC 300

Db 811 ATCTACTCAACTCAGAAATGAAGTAGCTGACAGCACTGGCTGAGGCTGAGTGCAGTGTTC 870

QY 301 GCTTGAAGGCGAGTCAGAAATGATGACTTCTGTGTGTATTGACCGCTGTGTGAACATG 360

Db 871 GCTTGAAGGCGAGTCAGAAATGATGACTTCTGTGTGTATTGACCGCTGTGTGAACATG 930

QY 361 GATGGGTGGACGCCAACATGATCTGTGATGATGGGGAGACTTAACCACTGGGTTAT 420

Db 931 GATGGGTGGACGCCAACATGATCTGTGATGATGGGGAGACTTAACCACTGGGTTAT 990

QY 421 AAGAATATCCAAACGTGTTAAGAGATCCAGGCATTTGTGAAGAGAGCGTACTGCT 480

Db 991 AAGAATATCCAAACGTGTTAAGAGATCCAGGCATTTGTGAAGAGAGCGTACTGCT 1050

QY 481 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTGTCCGCCATGAACGTC 540

Db 1051 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTGTCCGCCATGAACGTC 1110

QY 541 AATGATTCGTGTACCAACAGAAAGTTGATACTGTACTGCTGCGGAAATCCATTTTG 600

Db 1111 AATGATTCGTGTACCAACAGAAAGTTGATACTGTACTGCTGCGGAAATCCATTTTG 1170

QY 601 GATGGCTGAAGAGGACACAGATGTGATGTTGGTGGAAACAAGTGGTGTGTGGC 660

Db 1171 GATGGCTGAAGAGGACACAGATGTGATGTTGGTGGAAACAAGTGGTGTGTGGC 1230

QY 661 TATGTGAGGTAGGCAAGGGCTGTGCTCTCAAGCTCTTGAGCAATGTCTAC 720

Db 1231 TATGTGAGGTAGGCAAGGGCTGTGCTCTCAAGCTCTTGAGCAATGTCTAC 1290

QY 721 ATTACCGAAATGACCCCATCTGTGCTGTGCAAGGCTGCATGATGGGTTCAGGGTGCTA 780

Db 1291 ATTACCGAAATGACCCCATCTGTGCTGTGCAAGGCTGCATGATGGGTTCAGGGTGCTA 1350

QY 781 AAGCTAATGAAGTCATCCGGCAAGTCGATGCTGTAATACTTGACAGCAATAGAAT 840

Db 1351 AAGCTAATGAAGTCATCCGGCAAGTCGATGCTGTAATACTTGACAGCAATAGAAT 1410

QY 841 GTAGTGACAGGAGCACTTGATGCGATGAGAAACAGTTGTATGTCATATATGGGC 900

Db 1411 GTAGTGACAGGAGCACTTGATGCGATGAGAAACAGTTGTATGTCATATATGGGC 1470

QY 901 CACTCCAACACAGAAATCGATGTGACCAAGCTCCGCACTCCGAGCTGACGTGGAGCGA 960

Db 1471 CACTCCAACACAGAAATCGATGTGACCAAGCTCCGCACTCCGAGCTGACGTGGAGCGA 1530

QY 961 GTACGTTCTCAGGTGACCATGTCTATCTGGCCAGATGGCAACGAGTTGCTCTCTGGA 1020

Db 1531 GTACGTCTCAGGTGACCAATGTCATCTGGCCAGATGGCAAGAGTGTCTCTCTGGCA 1590
QY 1021 GAGGCTGCTCTACTCAATTTGAGCTGCTCCACAGTTCACCTTGTCTGTCCATCACA 1080
Db 1591 GAGGCTGCTCTACTCAATTTGAGCTGCTCCACAGTTCACCTTGTCTGTCCATCACA 1650
QY 1081 GCCACAACACAGGCTTTGGCACTGATAGACTCTATATGACCCGAGGGCGATACAA 1140
Db 1651 GCCACAACACAGGCTTTGGCACTGATAGACTCTATATGACCCGAGGGCGATACAA 1710
QY 1141 CAGGATGTGCT 1200
Db 1711 CAGGATGTGCT 1770
QY 1201 TCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGACTCAAC 1260
Db 1771 TCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGACTCAAC 1830
QY 1261 AAAAATGGCCATTCACCAACCTAATATTACAGATAC 1296
Db 1831 AAAAATGGCCATTCACCAACCTAATATTACAGATAC 1866

RESULT 6
AF315687 2677 bp mRNA linear PRI:22-MAR-2002
LOCUS AF315687
DEFINITION Homo sapiens S-adenosylhomocysteine hydrolase-like protein mRNA,
complete cds.
ACCESSION AF315687
VERSION AF315687.1 GI:16588686
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 2677)
AUTHORS Dekker, J.W., Budhia, S., Angel, N.Z., Cooper, B.J., Clark, G.J.,
Hart, D.N. and Kato, M.
TITLE Identification of an S-adenosylhomocysteine hydrolase-like
transcript induced during dendritic cell differentiation
JOURNAL Immunogenetics 53: (12) 993-1001 (2002)
MEDLINE 21901265
PUBMED 11904675
REFERENCE 2 (bases 1 to 2677)
AUTHORS Dekker, J.W., Budhia, S., Angel, N.Z., Cooper, B.J., Clark, G.J.,
Hart, D.N.J. and Kato, M.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-2000) Dendritic Cell Research, Mater Medical
Research Institute, Level 3, Aubigny Place, South Brisbane,
Queensland 4101, Australia

FEATURES
source
1. 2677
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p21"
/clone="211(1)B"
/cell_line="Hodgkin's disease cell line L428"
1. 114
/note="derived from 5' RACE"
369. 1961
/note="DCAL; dendritic cell expressed AHcy-like protein;
expressed in dendritic cells in blood, langerhans cells
and dermal dendritic cells; mRNA increased during monocyte
differentiation to monocyte-derived dendritic cells in the
presence of GM-CSF and IL-4"
/codon_start=1
/product="S-adenosylhomocysteine hydrolase-like protein"
/protein_id="AAL26869.1"
/db_xref="GI:16588687"
ADDMOFTKPKTKGRSLRSISQSDSYSAASYTDSDEVSPREKQGTNSKGS

SNFCVKNIKQAEFRREIEIAEQDMSALISLRKROGEKPLAGAKIVGCTHITAQTAV
LIEIICALGAOCRMSCANITYSTONENVAALAEAVAFAMKGESEDEFWCIDRCVNM
DGMAMMILDDGDLTHWVKKYPNVFKIRIGVIESVTGVHRLYQLSRAGKLCVPM
NNDVSVKOKFDNLCXCCRESILDLKRTDVMGKOVVGVGKGEVGVKALKALG
AIVYTEIDPICALQACMDGRVVKLVNVIROYDVITCTGNKNVYTRHIDRMKNSC
IVCNMGSNTEIDVYSLEPTELFMERVRSQVDHYIAPDGRVLLAEGRLLNLSCTV
PTEVLSTATQALALIELYNAPGRYKQDVYLLPRKMEYVASLHPSPDAHLTELT
DDQAKYLGINKNGPREPNVYR"
BASE COUNT 657 a 648 c 708 g 664 t
ORIGIN

Query Match 100.0%; Score 1296; DB 9; Length 2677;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCAAGGCGAGCAGCAATTTCTGTGTGAGAGAACATCAAGCAGGCAATTTGGACGCCG 60
Db 663 TCCAAGGCGAGCAGCAATTTCTGTGTGAGAGAACATCAAGCAGGCAATTTGGACGCCG 722
QY 61 GAGATTGAGATTGCAGAGCAAGACATGCTCTGATTTCACCTCAGGAAACGTGCTCAG 120
Db 723 GAGATTGAGATTGCAGAGCAAGACATGCTCTGATTTCACCTCAGGAAACGTGCTCAG 782

QY 121 GGGAGAGCCCTTGCTGCTGCTAAATAGTGGGCTGTACACACATCACAGCCAGACA 180
Db 783 GGGAGAGCCCTTGCTGCTGCTAAATAGTGGGCTGTACACACATCACAGCCAGACA 842
QY 181 GCGGTGTTGATTGAGACACTCTGTGCTGGGGGCTCAGTGCCGCTGCTGCTGTAA 240
Db 843 GCGGTGTTGATTGAGACACTCTGTGCTGGGGGCTCAGTGCCGCTGCTGCTGTAA 902

QY 241 ATCTACTCAACTCAGAAATGAGTAGCTGACGACTGGCTGAGGCTGAGGTTGCACTGTC 300
Db 903 ATCTACTCAACTCAGAAATGAGTAGCTGACGACTGGCTGAGGCTGAGGTTGCACTGTC 962
QY 301 GCTTGAAGGCGAGTCAAGAGATGACTCTGTGTGATTGACCGCTGTGTAACATG 360
Db 963 GCTTGAAGGCGAGTCAAGAGATGACTCTGTGTGATTGACCGCTGTGTAACATG 1022

QY 361 GATGGGTGGAGGCCCAACATGATCTCTGATGATGGGGAGACTTAACCACTGGGTTAT 420
Db 1023 GATGGGTGGAGGCCCAACATGATCTCTGATGATGGGGAGACTTAACCACTGGGTTAT 1082
QY 421 AAGAAGTATCCAAACGTGTTAAGAAGATCCAGGCAATTTGGAAGAGCGTGACTGCT 480
Db 1083 AAGAAGTATCCAAACGTGTTAAGAAGATCCAGGCAATTTGGAAGAGCGTGACTGCT 1142

QY 481 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTGTTCCGCCATGAACGTC 540
Db 1143 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTGTTCCGCCATGAACGTC 1202
QY 541 AATGATCTGTATCCAAACAGAGATTGTAATCTGTACTGCTGCCGGAATTCATTTTG 600
Db 1203 AATGATCTGTATCCAAACAGAGATTGTAATCTGTACTGCTGCCGGAATTCATTTTG 1262

QY 601 GATGGCTGAGAGAGACACACAGATGTGATGTTGGTGGGAACAAGTGGTGTGTCG 660
Db 1263 GATGGCTGAGAGAGACACACAGATGTGATGTTGGTGGGAACAAGTGGTGTGTCG 1322
QY 661 TATGGTGAAGGCAAGGCTGCTGTGCTCTCAAGCTCTTGAGCAATTTGCTTAC 720
Db 1323 TATGGTGAAGGCAAGGCTGCTGTGCTCTCAAGCTCTTGAGCAATTTGCTTAC 1382

QY 721 ATTACCGAAATGACCCCATCTGTGCTCTGAGGCTCATGATGGGTGAGGGTGTA 780
Db 1383 ATTACCGAAATGACCCCATCTGTGCTCTGAGGCTCATGATGGGTGAGGGTGTA 1442
QY 781 AAGCTAATGAGTATCCCGCAAGTGAATGCTAATACTTGACAGGAATAAGAAT 840
Db 1443 AAGCTAATGAGTATCCCGCAAGTGAATGCTAATACTTGACAGGAATAAGAAT 1502

QY 841 GTAGTACACGGGAGCACTTGATCCGATGAATAACAGTGTATCGTATCAATATGGGC 900
Db 1502 GTAGTACACGGGAGCACTTGATCCGATGAATAACAGTGTATCGTATCAATATGGGC 960

Db 1503 GTAGTGACACGAGACACTTGATCGCATGAAAAACAGTTGTATCGTATGCAATATGGGC 1562
QY 901 CACTCCACACAGAAATCGATGTGACACAGCCTCCGCACTCCGAGCTGACGTGGAGCGA 960
Db 1563 CACTCCACACAGAAATCGATGTGACACAGCCTCCGCACTCCGAGCTGACGTGGAGCGA 1622
QY 961 GTACGTTCTCAGGTGGACCATGTCTATCTGGCCAGATGGCAACGAGTGTCTCTCTGCGA 1020
Db 1623 GTACGTTCTCAGGTGGACCATGTCTATCTGGCCAGATGGCAACGAGTGTCTCTCTGCGA 1682
QY 1021 GAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTCCACCTTTGTTCTGTCCATCACA 1080
Db 1683 GAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTCCACCTTTGTTCTGTCCATCACA 1742
QY 1081 GCCACAACACAGGCTTTGGCACTGATAGAACTCTATTAATGCACCCGAGGGCGATACAG 1140
Db 1743 GCCACAACACAGGCTTTGGCACTGATAGAACTCTATTAATGCACCCGAGGGCGATACAG 1802
QY 1141 CAGGATGTGTACTTGTCTCTAAGAAATGATGAATACGTTGCCAGCTTGACATCTGCCA 1200
Db 1803 CAGGATGTGTACTTGTCTCTAAGAAATGATGAATACGTTGCCAGCTTGACATCTGCCA 1862
QY 1201 TCATTTGATGCCCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGAGCTCAAC 1260
Db 1863 TCATTTGATGCCCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGAGCTCAAC 1922
QY 1261 AAAAATGGGCATTCACAACTTAATATTACAGATAC 1296
Db 1923 AAAAATGGGCATTCACAACTTAATATTACAGATAC 1958

RESULT 7

HSMB00298 2510 bp mRNA linear PRI 18-FEB-2000
LOCUS Homo sapiens mRNA; cDNA DKFZp564A1523 (from clone DKFZp564A1523);
DEFINITION partial cds.

ACCESSION AL049954
VERSION AL049954.1 GI:4884203

KEYWORDS

SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE

1 (bases 1 to 2510)
Wambutt, R., Heubner, D., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
Direct Submission
Submitted (15-MAY-1999) MIPS, Am Klopferstutz 18a, D-82152
Martinsried, GERMANY

COMMENT

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
consortium of the German Genome Project.
This clone (DKFZp564A1523) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at <http://www.mips.biochem.mpg.de/proj/cDNA/>.

FEATURES

source
1..2510
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="1, 15.55 cR from DIS248"
/clone="DKFZp564A1523"
/tissue_type="brain"
/clone_1lb="564 (synonym: hfbz2). Vector pAMP1; host
X1-2blue; sites NotI + SalI"
/dev_stage="fetal"
1..1796
/gene="DKFZp564A1523"
<1..1796
/gene="DKFZp564A1523"
/note="S-adenosyl homocysteine hydrolase homolog XPYKona
splice variant"

/codon_start=3
/product="hypothetical protein"
/protein_id="CAB43223.1"
/db_xref="GI:4884204"
/translation="GLTVELDSSVCYILLLRHGSGRKGICAEISOPALPLPQE
TOAPSGESENRLNLRGRRAESONLSLRRLVSAQNGRSDWRMLTQDMEIDSLE
ISCFLOIQFADDMQEFKPTKTGRRLSRISQSSDTSYSSAAYTDSDEVPRE
KOQTSKGSNRCVKNIKQAEFRREIEADOMSAISLRKRAOGEKPLAGAKIVGC
THITAOAVLLETICALGAQCRWSACNISTONEVAALAEAGVAATPANKGESEDDFW
WCIDRCVNMDSQANMILDDGDLTHVYKKNPVEKIRGIVESESVYHRLYOLSK
AGKLCVPAMVNDVTKQKFDNLVCCRESILDLKRTDVMFGKQVYVCGYGEVKG
CCALKALGAIYITEIDPICALQACMDGRVYKLNKNEVIRQDVVITCGKNVYRE
HLDRMKNSCIVNMGHSNTEIDVLSLRTPELWERYRVSQVDHVWPGKRVLLAEGR
LNLSCSTVPFVLSITATQALALIELYNAPGRYKQDVYLLPKMDEVYASLHLP
FDAHLTELTDQAKYLLGNKNGPFPNYRY"
BASE COUNT 664 a 556 c 618 g 672 t
ORIGIN

Query Match 99.9%; Score 1294.4; DB 9; Length 2510;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1295; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCCTAAGGCGACGACCAATTTCTGTGTGAAGAATCAATCAAGCAGAGCAATTTGGACGCCG 60
Db 498 TCCTAAGGCGACGACCAATTTCTGTGTGAAGAATCAATCAAGCAGAGCAATTTGGACGCCG 557
QY 61 GAGATTGAGATTGACAGAGCAAGACATGCTGCTGTGATTTCATCAGGAACGTCGTCAG 120
Db 558 GAGATTGAGATTGACAGAGCAAGACATGCTGCTGTGATTTCATCAGGAACGTCGTCAG 617
QY 121 GGGAGAGACCCCTTGCTGCTGCTAAATAGTGGGCTGTACACATCAAGCCACAGACA 180
Db 618 GGGAGAGACCCCTTGCTGCTGCTAAATAGTGGGCTGTACACATCAAGCCACAGACA 677
QY 181 GCGGTGTGATTGAGACACTCTGTCCTGCGGCTGAGTGCCTGCTGTGTAAC 240
Db 678 GCGGTGTGATTGAGACACTCTGTCCTGCGGCTGAGTGCCTGCTGTGTAAC 737
QY 241 ATCTACTCACTCAGATGAAGTACGCTGACGACACTGGCTGAGGCTGAGTGCAGTGTTC 300
Db 738 ATCTACTCACTCAGATGAAGTACGCTGACGACACTGGCTGAGGCTGAGTGCAGTGTTC 797
QY 301 GCTTGAAGGGCGAGTCAAGATGACTTCTGTGTGTATGACCGCTGTGTAACATG 360
Db 798 GCTTGAAGGGCGAGTCAAGATGACTTCTGTGTGTATGACCGCTGTGTAACATG 857
QY 361 GATGGCTGACGAGCCACATGATGATGATGATGGGAGACCTTAACCACTGGGTTAT 420
Db 858 GATGGCTGACGAGCCACATGATGATGATGATGGGAGACCTTAACCACTGGGTTAT 917
QY 421 AAGAATATCCAAACGTGTTAAGAGATCCGAGCATGTGGAAGAGAGCGTGAATGCT 480
Db 918 AAGAATATCCAAACGTGTTAAGAGATCCGAGCATGTGGAAGAGAGCGTGAATGCT 977
QY 481 GTTCAAGGCTGTATCAGCTCTCCAAGCTGGGAAGCTGTGTTCGGGCATGAAGCTC 540
Db 978 GTTCAAGGCTGTATCAGCTCTCCAAGCTGGGAAGCTGTGTTCGGGCATGAAGCTC 1037
QY 541 AATGATCTGTTACCAACAGAGTTGATACTGTACTGTCTGCGGAGAAATCCATTG 600
Db 1038 AATGATCTGTTACCAACAGAGTTGATACTGTACTGTCTGCGGAGAAATCCATTG 1097
QY 601 GATGGCTGAAGAGGACACAGATGTGATGTTGGTGGGAACAGTGGTGTGGC 660
Db 1098 GATGGCTGAAGAGGACACAGATGTGATGTTGGTGGGAACAGTGGTGTGGC 1157
QY 661 TATGTGAGGTAGGCAAGGGCTGCTGCTCTCAAGCTCTGGAACAATTTGCTAC 720
Db 1158 TATGTGAGGTAGGCAAGGGCTGCTGCTCTCAAGCTCTGGAACAATTTGCTAC 1217
QY 721 ATTACGAAATGACCCCATCTGTCTCTGACAGGCTGCATGATGGGTTCAAGGTGTA 780
Db 1217 ATTACGAAATGACCCCATCTGTCTCTGACAGGCTGCATGATGGGTTCAAGGTGTA 780

Db 1218 ATTACGAAATCGACCCCATCTGCTCTGACAGGCTTCATGATGGGTTCAGGGTGTA 1277
QY 781 AAGCTAAATGAGTCAATCCGGCAGTGTGCTGTAATTAATTCAGACAGGAATAGAAT 840
Db 1278 AAGCTAAATGAGTCAATCCGGCAGTGTGCTGTAATTAATTCAGACAGGAATAGAAT 1337
QY 841 GTAGTACACGGGAGCAGTGTGATGCAATGAAACAGAGTGTGATGCAATATGGGC 900
Db 1338 GTAGTACACGGGAGCAGTGTGATGCAATGAAACAGAGTGTGATGCAATATGGGC 1397
QY 901 CACTCCACACACAGAAATCGATGTGACAGCCTCCGCACTCCGAGCTGACGTGGAGCGA 960
Db 1398 CACTCCACACACAGAAATCGATGTGACAGCCTCCGCACTCCGAGCTGACGTGGAGCGA 1457
QY 961 GTACGTTCTCAGGTGAGCAGTGTGCTGCGCAGATGGCAACGAGTTGCTCTCTGCGCA 1020
Db 1458 GTACGTTCTCAGGTGAGCAGTGTGCTGCGCAGATGGCAACGAGTTGCTCTCTGCGCA 1517
QY 1021 GAGGTCGCTCTACTCAATTTGAGCTGCTCCACAGTTCACACCTTGTCTGTCATCACA 1080
Db 1518 GAGGTCGCTCTACTCAATTTGAGCTGCTCCACAGTTCACACCTTGTCTGTCATCACA 1577
QY 1081 GCCACACACAGGCTTTGGCAGCTGATGAACTCTATATGACACCGGAGGGCGATACAAG 1140
Db 1578 GCCACACACAGGCTTTGGCAGCTGATGAACTCTATATGACACCGGAGGGCGATACAAG 1637
QY 1141 CAGGATGTACTTGTCTCCTAAGAAATGATGATGATGATGATGATGATGATGATGATG 1200
Db 1638 CAGGATGTACTTGTCTCCTAAGAAATGATGATGATGATGATGATGATGATGATGATG 1697
QY 1201 TCATTGATGCCCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAAC 1260
Db 1698 TCATTGATGCCCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAAC 1757
QY 1261 AAAAAATGGCCATTCACAACTTAATTAATACAGATAC 1296
Db 1758 AAAAAATGGCCATTCACAACTTAATTAATACAGATAC 1793
RESULT 8
LOCUS BC018218 1772 bp mRNA linear ROD 07-AUG-2002
DEFINITION Mus musculus, S-adenosylhomocysteine hydrolase-like 1, clone
ACCESSION BC018218 GI:17390492
VERSION BC018218.1
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1772)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (03-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Huiyik, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 23 Row: n Column: 6
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Genomescan gene prediction, Similarity but not identity
to protein.

FEATURES

source

1. 1772
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="CZECH II"
/clone="MGC:18748 IMAGE:4007102"
/tissue_type="Mammary tumor metastasized to lung."
MTV-LTR/mtl model. Expression driven by an MTV-LTR
enhancer."
/clone_1lb="NCI-CGAP_Lu30"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
191. 1693
/codon_start=1
/product="S-adenosylhomocysteine hydrolase-like 1"
/protein_id="AAH18218.1"
/db_xref="GI:17390493"
/db_xref="LocusID:10768"
/translation="MATVTAPKKQIOFADDMOEFKPTKTRSLRSISQSDTS
YSSAASYTSSDDEVSPREKQINSKSSNEFCVKNIKAEFRREIEIAEQMSALIS
LRRKQGEKPLAGAKIVGCTHTTAQAVLIETLALGACQWSACNISTONEVAAL
AEAGVAVFAWKGESEDEFWNCIDRCVNMDDQANMILDDGDLTHVYKYPNFKKI
RGIVESVTGVRHLYOLSKAKICVPMNVNDSVTKQRFNLYCCRESILDLKRTD
VMFGKQVNVCGYGEVKGCCALKALGAIIVITEIDPCALQACMDGFRVVKLNEVI
ROYDVITCTGNKNVTREHDMKNSCIYCNMGSNTEIDVSLRTPELTWERVRSQ
VDHVIWPDGKRVLLAEGRLNLSCSTVPTFLVSLTATQALALIELYNAPGRRKQD
VYLLPKMDEVASLHLPSEFAHLELTDQAKYLGKNNKPFKPNYRY"

CDS

BASE COUNT 462 a 406 c 506 g 398 t
ORIGIN

Query Match 90.0%; Score 1166.4; DB 10; Length 1772;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 1215; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 1 TCCAAGGGCAGCAGCAATTTCTGTGTAAGAAATCAAGCAGGAGCAATTTGGACGCCG 60
Db 395 TCGAAGGGCAGCAGCAATTTCTGTGTAAGAAATCAAGCAGGAGCAATTTGGACGCCG 454
QY 61 GAGATTGAGATTGAGCAAGACATGTCTGCTGATTTCACTCAGAAACGTGCTCAG 120
Db 455 GAGATTGAGATTGAGCAAGACATGTCTGCTGATTTCACTCAGAAACGTGCTCAG 514
QY 121 GGGGAGAGCCCTTGCTGCTGCTAAATAGTGGGCTGTACACACATCAGCCAGACA 180
Db 515 GGAGAGAGCCCTTGCTGCTGCTAAATAGTGGGCTGTACACACATCAGCCAGACA 574
QY 181 GCGGTGTGATTGAGACACTCTGTGCCCCCTGGGGCTCAGTGCCGCTGGTCTGTAAAC 240
Db 575 GCGGTATTATTGAGACCTTTGTGCCCTGGAGCTCAGTGCCGCTGGTCTGTAAAC 634
QY 241 ATCTACTCACTCAGAAATGAGTAGCTGACAGCACTGGCTGAGGCTGGAGTTGAGTTC 300
Db 635 ATCTATTCACACTCAGAAATGAGTAGCTGACAGCACTGGCTGAGGCTGGAGTTGAGTTC 694
QY 301 GCTTGAAGGGCGAGTCAAGAGATGACTTCTGTGTGTTATTGACCCGCTGTGTAACATG 360
Db 695 GCTTGAAGGGCGAGTCAAGAGATGACTTCTGTGTGTTATTGACCCGCTGTGTAACATG 754
QY 361 GATGGGTGGCAGGCCACATGATCCTGATGATGGGGAGAGACTTAACCCACTGGGTTAT 420
Db 755 GATGGGTGGCAGGCCATGATCCTGATGATGGGGAGAGACTTAACCCACTGGGTTAT 814
QY 421 AAGAAGTATCCAAACGTGTTAAGAGATCCGAGGCAATTTGGAGAGAGAGCGTGACTG 480
Db 815 AAGAAGTATCCAAACGTGTTAAGAGATCCGAGGCAATTTGGAGAGAGAGCGTGACTG 874

```

OY 481 GTTCACAGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTGTCCGCCATGAACGTC 540
    |||||||
DB 875 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTGTCCGCCATGAATGTC 934
OY 541 AATGATCTCTTACCAACAGAGTTGTAATCTGTACTGTGCGGCAATCCATTTG 600
    |||||||
DB 935 AATGATCTCTTACCAACAGAGTTGTAATCTGTACTGTGCGGCAATCCATTTG 994
OY 601 GATGGCCTGAAGAGACACAGATGTGATGTTGGTGGAAACAAGTGTGTGTGGC 660
    |||||||
DB 995 GATGGCCTGAAGAGACACAGATGTGATGTTGGTGGAAACAAGTGTGTGTGGC 1054
OY 661 TATGCTGAGTGAAGCAAGGCTGTGTGCTCTCTCAAGCTCTTGAGCAATGTCTAC 720
    |||||||
DB 1055 TATGCTGAGTGAAGCAAGGCTGTGTGCTCTCTCAAGGCTCTTGAGCAATGTCTAC 1114
OY 721 ATTACCGAATCGACCCCATCTGTGCTCTGAGAGCCCTGCATGATGGGTTCAGGGTGTA 780
    |||||||
DB 1115 ATCACAGAAATGACCCCATCTGTGCTCTGAGAGCCCTGCATGATGGGTTCAGGGTGTA 1174
OY 781 AAGCTAATGAGTGCATCCGGCAAGTGCATGCTAATACTTGACAGAAATAAGAAAT 840
    |||||||
DB 1175 AAGCTAATGAGTGCATCCGGCAGGTGAGCGTTGTAATTAATTGACAGAAATAAGAAAT 1234
OY 841 GTAGTGACAGGAGACACTTGATGCGATGAATAAACAAGTTGATGATGCAATATGGGC 900
    |||||||
DB 1235 GTAGTGACAGGAGACACTTGATGCGATGAATAAACAAGTTGATGATGCAATATGGGC 1294
OY 901 CACTCCAACACAGAAATGATGTGACAGGCTCCGCACTCCGAGCTGAGCTGGAGCGCA 960
    |||||||
DB 1295 CACTCCAACACAGGATGATGTGACAGGCTCCGCACTCCGAGCTGAGCTGGAGCGCT 1354
OY 961 GTACGTTCTCAGTGGACCATGTCTGCGGCAAGATGGCAACGAGTGTCTCTGCGCA 1020
    |||||||
DB 1355 GTACGTTCTCAGTGGACCATGTCTGCGGCAAGATGGCAACGAGTGTCTCTGCGCA 1414
OY 1021 GAGGCTGCTACTCAATTTGAGCTGTCCACAGTCCACCTTTGTTCTGCTCATCACA 1080
    |||||||
DB 1415 GAGGCTGCTACTCAATTTGAGCTGTCCACAGTCCACCTTTGTTCTGCTCATCACA 1474
OY 1081 GCCACAACACAGGCTTTGGCACTGATAGAACTCTAATAATGCACCCGAGGCGATACAAG 1140
    |||||||
DB 1475 GCTACAACACAGGCTTTGGCACTGATAGAACTCTAATAATGCACCCGAGGCGATACAAG 1534
OY 1141 CAGGATGTGTACTGCTTCTTAAGAAATGATGAATACGTTGCCAGCTTGCACTGCCA 1200
    |||||||
DB 1535 CAGGATGTGTACTGCTTCTTAAGAAATGATGAATATGTTGCCAGCTTGCACTGCCA 1594
OY 1201 TCATTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAAC 1260
    |||||||
DB 1595 TCATTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAAC 1654
OY 1261 AAAAATGGCCATCAAACTAATTATTACAGATAC 1296
    |||||||
DB 1655 AAAAATGGCCATCAAACTAATTATTACAGATAC 1690

```

RESULT 9

AK025372

LOCUS

AK025372

1987 bp

mRNA

linear

PRI 29-SEP-2000

DEFINITION

Homo sapiens CDNA: FLJ21719 fls, clone COLF0094.

ACCESSION

AK025372.1

GI:10437875

oligo capping; fls (full insert sequence).

Homo sapiens colon mucosa CDNA to mRNA, clone_11b:COLF

clone:COLF0094.

SOURCE

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 (sites)

Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,

Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,

Nakamura,Y., Isogai,T. and Sugano,S.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 (sites)

Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,

Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,

Nakamura,Y., Isogai,T. and Sugano,S.

AUTHORS

Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,

Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,

Nakamura,Y., Isogai,T. and Sugano,S.

```

TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1987)
AUTHORS Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
        Shibahara,T., Tanaka,T. and Nakamura,Y.
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
        University of Tokyo, Laboratory of Genome Structure Analysis, Human
        Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
        Japan (E-mail:cdna@leims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
        Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
        International Trade and Industry of Japan; cDNA full insert
        sequencing: Research Association for Biotechnology; cDNA library
        construction, 5'- & 3'-end one pass sequencing: Department of
        Virology and Human Genome Center, Institute of Medical Science,
        University of Tokyo (partly supported by science and Technology
        Agency).
FEATURES
    source      1. 1987
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="COLF0094"
                /tissue_type="colon mucosa"
                /clone_lib="ColF"
                /note="cloning vector pME18SFL3"
BASE COUNT 557 a 430 c 501 g 499 t
ORIGIN

```

```

Query Match 66.6%; Score 863; DB 9; Length 1987;
Best Local Similarity 79.2%; Pred. No. 3.8e-258;
Matches 1025; Conservative 0; Mismatches 270; Indels 0; Gaps 0;

```

```

OY 1 TCCAGAGGACAGCAATTTCTGTGTGAAGAACATCAAGCAGGAGATTTGGAGCCCG 60
    |||||||
DB 500 TCTAAGGAGACAGTACTCTGTGTGAAGAACATCAAGCAGGAGATTTGGAGAGAGA 559
OY 61 GAGATTGAGATGACAGCAAGACATGTCTGCTGATTTCACTCAGAAACCTGCTCAG 120
    |||||||
DB 560 GAAATTGAATTTGCTGAACAAGAAATGCTGCTGATGAGCTTGAAGAGAGAGCTCAA 619
OY 121 GGGAGAGAGCCCTTGCTGTGCTAAATAGTGGCTGTACACACATCAGAGCCAGACA 180
    |||||||
DB 620 GGAGAAAGCCCTTGCTGTGAGCCAAATGCTGGTGTGACACACATCAGTGTGAGACT 679
OY 181 GCGGTGTGATGAGACACTCTGTGCCCCGAGGCTCAGTCCGCTGCTGTGTAAC 240
    |||||||
DB 680 GCTGTGCTTATGAACACTCTGGGTGCTGTGGGGCCAGTGGCGTGGCTGTGCAAC 739
OY 241 ATCTACTCAACTCAGATGAAGTAGCTGCAGCACTGGCTGAGGCTGAGTGTGACTC 300
    |||||||
DB 740 ATCTATTCACCTCTCAATGAAGTGGCTGCTGCTTACAGAAAGTGGATTTCTGTTT 799
OY 301 GCTTGAAGGGGAGTCAAGATGACTTCTGTGTGTATTTGACCGCTGTGTGAACATG 360
    |||||||
DB 800 GCCTGGAAGGGAGTCAAGATGACTTCTGTGTGTATGATAGATGTGTGAATGTG 859
OY 361 GATGGTGGCAGGCAACATGATCTGTGATGATGGGAGAGCTTAACCCACTGGTTAT 420
    |||||||
DB 860 GAGGCTGGCAGGCAACATGATCTGTGATGATGGGAGAGCTTAACCCACTGGTTAT 919
OY 421 AAGAGTATCCAAAGTGTTTAAGAGATCCGAGCATTTGTGAAGAGAGCGTGAATG 480
    |||||||
DB 920 AAAAGTATCCCAACATGTTTAAGAAATCAAGGCGATAGTAGAGAGAGTGTACTGA 979
OY 481 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTCTCCGGCCATGAAGTC 540
    |||||||
DB 980 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTCTCCAGCCATGAATGTC 1039
OY 541 AATGATCTCTTACCAACAGAGTTGATTAATCTGTACTGTGCTGCCAGAGATCCATTTG 600
    |||||||
DB 1040 AATGATCTCTTACCAACAGAGTTGATTAATCTGTACTGTGCTGCCAGAGATCCATTTG 1099

```

```

QY 601 GATGGCCTGAGAGACACAGATGTGTGTTGGTGGGAACAAGTGTGTGTGGC 660
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1100 GATGCACTTAAAGAGACACAGATGTGTGTTGGTGGGAACAAGTGTGTGGC 1159
QY 661 TATGGTAGAGTAGGCAAGGCTGTGTGTCTCTCAAGCTCTTGAGCAATGTCTAC 720
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1160 TATGAGAGAGTGGGGAAGGCTGTGTGTCTCTCAAGCTCTTGAGCAATGTCTAC 1219
QY 721 ATTACGAAATGACCCCACTGTGTCTCTGAGGCTGTGATGGTTCAGGCTGTGTA 780
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1220 GTAACGAAATGACCCCACTGTGTCTCTGAGGCTGTGATGGTTCAGGCTGTGTA 1279
QY 781 AAGCTAATGAGTCACTCCGCAAGTGCATGTCTAATACTTGACAGGAATAAGAAT 840
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1280 AATTAATGAGTCACTCCGCAAGTGCATGTCTAATACTTGACAGGAATAAGAAT 1339
QY 841 GTAGTGACACGGGACACTGTGATCGCATGAAAAACAGTGTATCGTATGCAATATGGC 900
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1340 GTGATTAACAGAGAGACACTGTGACCGTATGAAGAATAGCTGCATCGTTTGAACATGGGA 1399
QY 901 CACTCCACACAGAAATCGATGTGACCAAGCTCTCCGCACTCCGAGCTGACGTGGAGCGA 960
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1400 CATTCCACACAGAGATGTGACGTGGCGAGTCTGCGGACACAGAACTGACCTGGAGCGA 1459
QY 961 GTACGTTCTCAGTGGACCATGTCTGTGCGCAGATGGCAACAGAGTGTCTCTCTGCGCA 1020
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1460 GTGAGATCTCAAGTGTGACCATGTGATGGCCTGTGATGGCAAGAGATGACTGTGCGCA 1519
QY 1021 GAGGCTGCTACTCAATTTGAGCTGTCCACAGTTCACACCTTTGTCTGTCTCATCACA 1080
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1520 GAGGGCCGCTGTGTAACCTGTGCTCCACAGTGTGCTCATATTTGTCTCATCACA 1579
QY 1081 GCCACAACACAGGCTTTGGCACATGATAGAACTCTAATAATGACACCCGAGGGCGATACAAG 1140
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1580 GCTACTACTCAGGCTCTTGCCCTGTGATAGAGCTTTACAATGCTCTGAGGGCTGCTATAAG 1639
QY 1141 CAGGATGTGTACTGTCTCTTAAGAAAAATGATGAATACGTTGCCAGCTTGCACTGCCA 1200
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1640 CAGGATGTGTACTGTCTCTTAAGAAAAATGATGAATACGTTGCCAGCTTGCACTGCCA 1699
QY 1201 TCATTTGATGCCCACTTACAGAGCTGACAGATGACCAAGAAAAATATCTGGGACTCAAC 1260
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1700 ACCTTTGATGCCCACTTACAGAGCTGACAGATGACCAAGCAAGTATCTGGGACTCAAC 1759
QY 1261 AAAAATGGCCATTCAAACTAATATTATACAGATA 1295
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1760 AAGAATGGCCCTTCAAGCTTAATATTATACAGATA 1794

```

```

RESULT 10
BC008349 2030 bp mRNA linear PRI 12-JUL-2001
LOCUS Homo sapiens, similar to S-adenosylhomocysteine hydrolase-like 1,
DEFINITION clone IMAGE:3536052, mRNA, partial cds.
ACCESSION BC008349
VERSION BC008349.1 GI:14249335
KEYWORDS Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1; (bases 1 to 2030)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (25-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

```

```

REMARK NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTP/DTP
cDNA Library Preparation: Rubin Laboratory

```

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
 Susana Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
 Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
 Ness, Pawan Pandoh, Anna-Lisa Prabhoo, Parvaneh Saeedi, Jacqueline
 Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott,
 Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
 George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAL Plate: 23 Row: 3 Column: 18.

FEATURES

```

source
1..2030
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3536052"
/tissue_type="Lung, small cell carcinoma"
/clone_1ib="NIH_MGC_7"
/lab_host="DH10B-R"
/label="Vector: pOTB7"
<1..1838
/codon_start=3
/product="Similar to S-adenosylhomocysteine hydrolase-like
1"
/protein_id="AAH08349.1"
/db_xref="GI:14249336"
/translation="MSVQVSAASAAAKVPEVELKDISPSEASQGLSTAAGAMAP
AGGDEAPAPAPAEAPVPGVPGSAPAAALSPAGAPVQASAMKRSDDPHOHORHRRGG
EALVSPDGTVEAPRTVKQIQFADQKQEFNKRPRKIGRSLRSISQSTDSYSSAA
SYTSSDDETSPPROKROKNSGSSDFCVKNIKQAEGRREIEIAEQEMPALMLKRA
QGEKPLAGAKIVGCTHITAOAVLMEITGALGACRMAACNISTLNEVAALAESGF
PVFAWKGESEDEFWNCIDRCVNEGQPMILDDGDLTHWIKKYPNMEFKIKGIVE
ESVTGVHRLYQLSKAGRLCVPAMVNDVTKQKFDNLCCRESILDGLKRTDMEGG
KQVAVCGYGEVKGCCALAKAMGSIYVTEIDPICALQACMDGFRVLKLVYROYDI
VITCTGKNVNTREHIDRMKNSCIVCMGHSNTEIDVASLRTPELTWERVRSVDHVI.
WPDGKRVILAEGRILNLSCTVPTFVLSITATQALALIELYNAPEGRYKQDYLLP
KRMDEVASLHLPTFDHLELTLEDOAKYILGNKNGPFXPNYRY"

```

```

BASE COUNT 539 a 492 c 548 g 451 t
ORIGIN

```

```

Query Match 66.5%; Score 861.4; DB 9; Length 2030;
Best Local Similarity 79.1%; Pred. No. 1.2e-257;
Matches 1024; Conservative 0; Mismatches 271; Indels 0; Gaps 0;

```

```

QY 1 TCCAAGGCGACAGCAATTTCTGTGTGAGAACATCAAGCAGGCAATTTGACGCCG 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 540 TCTAAGGAGCAGTGAATCTGTGTGTAAGAACATCAAGCAGGCAATTTGACGAGA 599
QY 61 GAGATTGAGATTGACAGCAAGACATGCTGTGATTTCACTCAGGAAGAGTGTCTAG 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 600 GAAATTGAAATTGCTGACACAGAAATGCTGATGATGAGCTTTGAGGAAGAGCTCAA 659
QY 121 GGGAGAGCCCTTGCTGTGTGTAATAAGTGGGCTGTACACACATCACAGCCAGACA 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 660 GGAGAAAGCCTTTGGCTGTGAGCCAAATCGTGGGTGACACACATCACTGCTCAGACT 719
QY 181 GCGGTGTGATTGAGACACTGTGCGCTGGGGGCTCAGTGCCGCTGTGCTGTAAAC 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 720 GCTGTCTTATGGAACACTGTGGGTCTGTGGGGCCCAAGTGCCGATGGGCTGCTGCAAC 779
QY 241 ATCTACTCAACTCAGATGAGTAGAGTGCACACTGGCTGAGGCTGAGTGCAGTGTTC 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 780 ATCTATTCACCTCAATGAAAGTGCTGTCTGTAGCAGAAAGTGAATTTCTGTTT 839
QY 301 GCTTGAAGGCGAGTCAAGAGATGACTTCTGTGTGTATGACCGCTGTGGAACATG 360
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 840 GCCTGAAGGAGAGTCAAGAGATGACTTGTGTGTGTATGATGATGATGTGTAATGTG 899

```

QY	361	GATGGGTGGCAGGCCCAACATGATCCCTGGATGATGGGGAGACTTAAACCACCTGGGTTTAT	420
Db	900	GAGGGCTGGCAGGCCAAACATGATCTTGGATGATGAGGGGATCTTACCACCTGGATTAT	959
QY	421	AAGAAGTATCCAAACGTGTTTAAGAAGATCCGAGCATTGTGGAAGAGAGCGTGAAGT	480
Db	960	AAAAAGTATCCCAACATGTTTAAGAAATCAAGGGCATTAGAGAGAGAGTGTACTGCA	1019
QY	481	GTTACAGCGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTCTCCGGCCATGAAGTC	540
Db	1020	GTTACAGCGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTCTCCAGCCATGAATGTC	1079
QY	541	AATGATCTGTATCCAAACAGAAAGTTGATTAACCTTGTACTGTCTGCCGAAATCCATTG	600
Db	1080	AATGATCTGTATCCAAACAGAAATTTGACAAACCTGTACTGTCTGCCGAAATCCATTG	1139
QY	601	GATGGCCTGAAGAGGAGCACCAGATGTGATGTTGGTGGGAAACAAGTGGTGTGTGGC	660
Db	1140	GATGGCCTGAAGAGGAGCACCAGATGTGATGTTGGTGGGAAACAAGTGGTGTGTGGC	1199
QY	661	TATGGTGAAGTGAAGGAGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	720
Db	1200	TATGGTGAAGTGAAGGAGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1259
QY	721	ATTACCGAATCGACCCCATCTGTGCTCTGCAGGCGCTGCATGGATGGGTTCAAGGTGTA	780
Db	1260	GTAACCTGAATGACCCCATCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1319
QY	781	AAGCTAAATGAAGTCAATCCGCGCAAGTGCATGTCGTAATTAATTCACACAGGAATAGAAT	840
Db	1320	AAATTAATGAAGTCAATCCGCGCAAGTGCATGTCGTAATTAATTCACACAGGAATAGAAT	1379
QY	841	GTAAGTGAACAGGAGCAGCTTGGATGCGATGAAAAACAGTTGTATCGTATGCAATATGGC	900
Db	1380	GTAAGTGAACAGGAGCAGCTTGGATGCGATGAAAAACAGTTGTATCGTATGCAATATGGC	1439
QY	901	CATCTCAACACAGAAATGATGTGACACAGCCTCCGACCTCCGAGAGCTGACGTGGAGCGA	960
Db	1440	CATCTCAACACAGAAATGATGTGACACAGCCTCCGACCTCCGAGAGCTGACGTGGAGCGA	1499
QY	961	GTAAGTGAACAGGAGCAGCTTGGATGCGATGAAAAACAGTTGTATCGTATGCAATATGGC	1020
Db	1500	GTAAGTGAACAGGAGCAGCTTGGATGCGATGAAAAACAGTTGTATCGTATGCAATATGGC	1559
QY	1021	GAGGCTGCTACTCAATTTGAGCTGCTCCACAGTTCCACCTTTGTTCTGTCCATCACA	1080
Db	1560	GAGGCTGCTACTCAATTTGAGCTGCTCCACAGTTCCACCTTTGTTCTGTCCATCACA	1619
QY	1081	GCCACACACAGGCTTTGGCAGCTGATAGAAGTCTAATATGCAACCCGAGGGGATACAG	1140
Db	1620	GCTACTACTCAGGCTCTTGCTGCTGATAGAGCTTACAACTGCTCTGAGGGTCTATPAG	1679
QY	1141	CAGGATGTGACTTGTCTCTTAAGAAATGATGAATACGTTGCCAGCTTGCATCTGCCA	1200
Db	1680	CAGGATGTGACTTGTCTCTTAAGAAATGATGAATACGTTGCCAGCTTGCATCTGCCA	1739
QY	1201	TCATTTGATGCCCCACCTTACAGAGCTGACAGATGACCAACAAATATCTGGGACTCAAC	1260
Db	1740	ACCTTTGATGCCCCACCTTACAGAGCTGACAGATGACCAACAAATATCTGGGACTCAAC	1799
QY	1261	AAAAATGGGCATTCAAACCTAATTATTACAGATA	1295
Db	1800	AAGAAATGGGCCTTCAAGCCTAATTACTACAGGTA	1834
RESULT 11			
AB020635			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			

ORGANISM	REFERENCE	TITLE	JOURNAL	MEDLINE	AUTHORS	TITLE	JOURNAL	FEATURES	source
SK plus clone:hn04230.	Homo sapiens								
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.									
1 (sites)									
Nagase,T., Ishikawa,K., Suyama,M., Kikuno,R., Hiroseawa,M.,									
Miyajima,N., Tanaka,A., Kotani,H., Nomura,N. and Ohara,O.									
Prediction of the coding sequences of unidentified human genes.									
XII. The complete sequences of 100 new cDNA clones from brain which									
code for large proteins in vitro									
DNA Res. 5 (6), 355-364 (1998)									
99156230									
2 (bases 1 to 5025)									
Ohara,O., Suyama,M., Kikuno,R., Nagase,T. and Ishikawa,K.									
Direct Submission									
Submitted (02-DEC-1998) Osamu Ohara, Kazusa DNA Research Institute,									
Laboratory of DNA Technology; Yana 1532-3, Kisarazu, Chiba									
292-0812, Japan (E-mail:cdna@info.kazusa.or.jp, Tel:+81-438-52-3913,									
Fax:+81-438-52-3914)									
Location/Qualifiers									
1. 5025									
/organism="Homo sapiens"									
/db_xref="taxon:9606"									
/clone="hn04230"									
/sex="male"									
/tissue_type="brain"									
/clone_1b="pbluescriptII SK plus"									
/dev_stage="adult"									
1. 5025									
/gene="KIAA0828"									
<1. 1860									
/gene="KIAA0828"									
/codon_start=1									
/product="KIAA0828 protein"									
/protein_id="BAA74851.1"									
/db_xref="GI:4240145"									
/translation="EPVVAEEAVMSVOVSAASAAKVPVELKLDSPSEASOLGLST									
AAGMAPAPGAGGDEAPAPAPAEPRPVPGSGPAAALSPAAGKVAQASAMKRSDPHH									
QHORHRDGEALVSPDGTVEAPRTVKIQFADQKOEFNKRPRTIGRSLRSISOSS									
TDSYSSAASITDSSDDDETSPRDKQKNSGSSDCEVKNIKQAEFGREIEIAEQEMPA									
LMALRKRAQGEKPLAGAKIVGCTHITPAQAVIMETIGLAQCRMACNITSTLNEVA									
AALAESGEPVFAWKGESEDDFWWCIDRCYVVEGWQPMMLDDGDLTHIYKYPNMF									
KKIKGIVESVTGVRHLXOLSKAGKLCVPMANVNDSTVKQFQNDLCCRESILDGLKR									
TTDMFEGGKQVYVCGYGEVGGKCCALAKMGSIVYTEIDPICALQACMDGFRVLKLN									
EVIRQVDIVITCTGNKNVTRREHLDRMKNSCIVCNMGHSNTEIDVASIRPELTWERV									
RSQVDHIVPDGKRIIVLAEGRLLNLSCTVPTFVLSITATQALALIELYNAPEGRY									
KQDYYLLPKKMEYVASIHLPTFDALHTELDEQAKYLGINKNGPRFNNYRR"									
BASE COUNT	1281 a	1179 c	1136 g	1429 t					
ORIGIN									
Query Match	66.5%	Score 861.4;	DB 9;	Length 5025;					
Best Local Similarity	79.1%	Pred. No. 1.4e-257;							
Matches 1024;	Conservative 0;	Mismatches 271;	Indels 0;	Gaps 0;					
1	TCCAAGGGCAGCAGCAATTTCTGTGTGAAGAACAATCAAGCAGGCGAATTGGAGCGCG	60							
2	TCCTAAGGAGCAGCAGTGTCTGTGTGAAGAACAATCAAGCAGGCGAATTGGAGCGAGA	621				</			

OY 301 GCTGGAAGGGGAGTCAAGAGATGACTTCTGCTGTGTATGTACCCGCTGTGCAACATG 360
DB 862 GCCTGGAAGGGGAGTCAAGAGATGACTTCTGCTGTGTATGTACCATGATGTGTGAATGTG 921
OY 361 GATGGGTGGGAGCCCAACATGATCTGTGATGTGGGGAGACTTAACCCACTGGGTTAT 420
DB 922 GAGGGCTGGGAGCCCAACATGATCTGTGATGTGGGGAGACTTAACCCACTGGGTTAT 981
OY 421 AAGAAGTATCCAAACGTGTTTGAAGAGATCGAGGCATTTGTGAAGAGAGCGGTGACTGCT 480
DB 982 AAAAAGTATCCCAACATGTTTGAAGAAATCAAGGCATAGTAGAGAGAGTGTACTGGA 1041
OY 481 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTGTCCGGCATGAACGTC 540
DB 1042 GTTCACAGGCTGTATCAGCACTGTCCAAAGCTGGGAAGCTGTGTCCAGCCATGAATGTC 1101
OY 541 AATGATCTGTGTACCAACAGAAAGTTGATTAATCTGTACTGTCTGCGGAGAAATTCATTTTG 600
DB 1102 AATGATCTGTGTACCAACAGAAATTTGACAACTCTACTGTGCGGAGAAATTCATTTCTT 1161
OY 601 GATGGCTGGAAGAGACACAGATGTGATGTTTGTGGGAAACAAGTGTGTGTGTCGTC 660
DB 1162 GATGGACTTAAAGGACACACAGATGTGATGTTTGTGGGAAACAAGTGTGTGTGTCGTC 1221
OY 661 TATGGTGAAGTGAAGGAGGAGGCTGTGTGTCTCTCAAGCTCTTGAGAGCAATGTCTAC 720
DB 1222 TATGGAGAGGTGGGGAAGGAGGCTGTGTGTCTCTCAAGCTCTTGAGAGCAATGTCTAC 1281
OY 721 ATTACCGAAATCGACCCCATCTGTGTCTGAGGAGGCTGTGATGGGTTCAAGGTTGTA 780
DB 1282 GTAACTGAATGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1341
OY 781 AAGCTAAATGAAGTCAATCCGGAAGTGTGATTAATTAATTAATTAATTAATTAATTA 840
DB 1342 AAATTAATGAAGTCAATCCGGAAGTGTGATTAATTAATTAATTAATTAATTAATTA 1401
OY 841 GTAGTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900
DB 1402 GTGTAACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1461
OY 901 CACTCCAACACAGAAATGATGTGACCAAGCTCCGCACTCCGAGCTGAGCTGGGAGCGA 960
DB 1462 CATTCACACAGAGATGTGACGTCGCGAGTCTGCGGACACAGAACTGACCTGGGAGCGA 1521
OY 961 GTACGTCTCAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020
DB 1522 GTGAGATCTCAAGTGTGACCATGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1581
OY 1021 GAGGTCGTCTACTCAATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1080
DB 1582 GAGGTCGTCTACTCAATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1641
OY 1081 GCCACAACACAGGCTTTGGCAGTGAAGTCTTATATGACACCGGAGGAGGAGGAGGAGG 1140
DB 1642 GCTACTACTCAGGCTCTGCTGTGATGAGCTTTACAATGCTCCTGAGGAGTGGCTATAAG 1701
OY 1141 CAGGATGTACTTGTCTCTAAGAAATGATGAATGATGATGATGATGATGATGATGATG 1200
DB 1702 CAGGATGTACTTGTCTCTAAGAAATGATGAATGATGATGATGATGATGATGATGATG 1761
OY 1201 TCAATTTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
DB 1762 ACCTTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1821
OY 1261 AAAAATGGGCAATTCAAACCTTAATTAATTAATTAATTAATTAATTAATTAATTA 1295
DB 1822 AAGAATGGGCAATTCAAACCTTAATTAATTAATTAATTAATTAATTAATTAATTA 1856

RESULT 12
BC024325
LOCUS BC024325 5052 bp MRNA linear PRI 12-MAR-2002

DEFINITION Homo sapiens, KIA0828 protein, clone MGC:21525 IMAGE:3907552,
mRNA, complete cds.
ACCESSION BC024325
VERSION BC024325.1 GI:19353091
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 5052)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcdéparill.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAC Plate: 22 Row: 1 Column: 22
This clone was selected for full length sequencing because it
passed the following selection criteria: similarity but not
identity to protein.

FEATURES

source
1..5052
/organism="Homo sapiens"
/db_xref="locusID:23382"
/db_xref="taxon:9606"
/clone="MGC:21525 IMAGE:3907552"
/tissue_type="uterus, leiomyosarcoma"
/clone_lib="NIH_MGC_71"
/lab_host="DH10B"
/note="Vector: pCMV-Sport6"
55..1890
/codon_start=1
/product="KIA0828 protein"
/protein_id="AAH24325.1"
/db_xref="GI:19353092"
/translation="MSQVVSAAAKVPEVELKDLSPSEASQGLSTAAGAMAP
AGGDEPAPAPAEPRVPVPGSPFAAALSPAGKVPQASAKRSDPHQHRHDDG
EALVSPDGVTEAPRTVKQIQPADQKQENKRPKIGRSISRSIOSSTDSYSA
SYTSSDDETSPRDKQKNSGSDFCVKNIKQAEFGRELEIAEOMALALRKRA
OGERPLAGAKIVGCTHTAQRVILMETLGAOCRWAAACNTSYLNEVAALAESGF
PYPAWKGESEDDFWWCIDRCVAVEGWQPNMILDDGDLTHWLYKKYPNMEKKIKGIVE
ESYTGVRHLXQLSKAGKLCVPMANVNDVTKQKFDNLCCRESILDLKRTDMFEG
KQVYVCGYGEVKGCCALAKAGSIVYVTEIDPICALQACDGFRLVKNLVIQVDI
VITCTGNKNVTRHLDKRNKSCIVCMGHSNTEIDVASLRTPELTWERVRSQVDHI
WPDGRIVLLAEGRLNLSCSTVPTFVLSITATQALALIELYNAPREGRYKODVLLP
KMDIEVVASLHLPFDALHLELDEQAKYLGINKNGPFKNRYRYR"

CDS
BASE COUNT 1300 a 1184 c 1145 g 1423 t
ORIGIN

Query Match 66.5%; Score 861.4; DB 9; Length 5052;
Best Local Similarity 79.1%; Pred. NO. 1.4e-257;
Matches 1024; Conservative 0; Mismatches 271; Indels 0; Gaps 0;

OY 1 TCCAAGGGCAGACGAATTTCTGTGTGAAGACATCAAGCAGGAGATTGGAGCGCGG 60
DB 592 TCTAAGGGAAGAGTACTTGTGTGAAGACATCAAAACAGCAGAGTTGGACGAGA 651
OY 61 GAGATTGAGATTGACAGACATGCTGTCTGATTTCACTCAGAAACGTGCTCAG 120

```
Db 652 GAAATGAAATGCTGACACAGAAATGCTGATTTGATGGCTTTGAGGAAGAGCTCAA 711
Oy 121 GGGAGAGCCCTTGGCTGTGCTAAATAGTGGCTGTACACATCACAGCCAGACA 180
Db 712 GGAGAAAGCCTTGGCTGAGCCAAATCGTGGGTGACACACATCACTGCTCAGACT 771
Oy 181 GCGGTGATGATGAGACACTCTGCTCCCTGGGGGCTCAGTGCCTGCTGCTGTAC 240
Db 772 GCTGTGCTTATGGAACCTCTGGGTGCTCTGGGGGCCAGTCCGATGGGCTGCTGCAAC 831
Oy 241 ATCTACTCACTCAGAAATGAGTACGTCAGCAGCTGGCTGAGGCTGGAGTTGCACTGTC 300
Db 832 ATCTATTCACCTCTCAATGAGTGGCTGCTGCTCTAGCAGAAAGTGGATTTCCTGTTT 891
Oy 301 GCTTGAAGGGCGAGTCAAGAGTACTTCTGCTGTATGACCCGCTGTGTGAACATG 360
Db 892 GCTTGAAGGAGAGTCAAGAGTACTTCTGCTGTATGATGATGATGATGATGATGATG 951
Oy 361 GATGGTGCAGGCCAACATGATCTGATGATGATGATGATGATGATGATGATGATGATG 420
Db 952 GAGGCTGCAGCCAAACATGATCTGATGATGATGATGATGATGATGATGATGATGATG 1011
Oy 421 AAGAGTATCCAAACGTGTTAAGAGATCCAGGCACTGTGAGAGAGAGCGTACTGT 480
Db 1012 AAAAGTATCCCAACATGTTAAGAGATCCAGGCACTGTGAGAGAGAGCGTACTGT 1071
Oy 481 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTCTCCGCCATGAACGTC 540
Db 1072 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTCTCCGCCATGAATGTC 1131
Oy 541 AATGATCTGCTTACCAACAGAGTGTGATTAATCTGTACTGCTGCGGAGAAATCCATTG 600
Db 1132 AATGATCTGCTTACCAACAGAGTGTGATTAATCTGTACTGCTGCGGAGAAATCCATTG 1191
Oy 601 GATGGCTGAGAGAGACACAGATGTGATGATGATGATGATGATGATGATGATGATGATG 660
Db 1192 GATGGCTGAGAGAGACACAGATGTGATGATGATGATGATGATGATGATGATGATGATG 1251
Oy 661 TATGATGAGTGAAGGAGGCTGTGCTGCTCTCAAGCTCTTGAGCAATGTCTAC 720
Db 1252 TATGATGAGTGAAGGAGGCTGTGCTGCTCTCAAGCTCTTGAGCAATGTCTAC 1311
Oy 721 ATTACGAATTCAGCCCATCTGCTGCTGCAAGGCTGCATGATGGTTCAAGGCTGTA 780
Db 1312 GATGCTGAATTCAGCCCATCTGCTGCTGCAAGGCTGCATGATGGTTCAAGGCTGTA 1371
Oy 781 AAGCTAATGATGATCCCGGCAAGTGTGCTGATTAATCTGCAACAGAAATAGAAT 840
Db 1372 AATTAATGATGATCCCGGCAAGTGTGCTGATTAATCTGCAACAGAAATAGAAT 1431
Oy 841 GTAGTGACAGGAGCACTGTGATGCAAAAAACAGTGTATGCAATATGGGC 900
Db 1432 GTAGTGACAGGAGCACTGTGATGCAAAAAACAGTGTATGCAATATGGGC 1491
Oy 901 CACTCCAACAGAAATGATGTGACCACTCCGCACTCCGAGCTGAGCTGGAGCGA 960
Db 1492 CACTCCAACAGAAATGATGTGACCACTCCGCACTCCGAGCTGAGCTGGAGCGA 1551
Oy 961 GTAGCTTCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1020
Db 1552 GTAGCTTCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1611
Oy 1021 GAGGCTGCTACTCAATTTGAGCTGCTCCACAGTTCACACTTTGCTGCTCATCACA 1080
Db 1612 GAGGCTGCTACTCAATTTGAGCTGCTCCACAGTTCACACTTTGCTGCTCATCACA 1671
Oy 1081 GCCACACACAGGCTTTGGCACTGATAGACTCTATATGCAACCCGAGGGCGATACAG 1140
Db 1672 GCTACTACTCAGGCTCTTGCTGATAGAGCTTTACAATGCTCCTGAGGGTCTATTAAG 1731
Oy 1141 CAGGATGTGATGCTCTCTTAAGAAATGATGATGATGATGATGATGATGATGATGATG 1200
```

```
Db 1732 CAGATGTCTACTGCTGTCGCCAAGAGATGATGATGATGATGATGATGATGATGATG 1791
Oy 1201 TCATTTGATGCCCCACTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAC 1260
Db 1792 ACCTTTGATGCCCCACTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAC 1851
Oy 1261 AAAATGGCCATTCAAACCTAATTTATACAGATA 1295
Db 1852 AAGATGGCCCTTCAAGCCTAATTTACTACAGGTA 1886
```

```
RESULT 13
AF035319 1369 bp mRNA linear PRI 07-JUL-2000
LOCUS AF035319 Homo sapiens clone 23931 mRNA, partial cds.
DEFINITION AF035319
ACCESSION AF035319 GI:2661082
VERSION AF035319.1
KEYWORDS FLI-CDNA.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1369)
AUTHORS Andersson, B., Wentland, M.A., Ricafrente, J.Y., Liu, W., and Gibbs, R.A.
TITLE A 'double adaptor' method for improved shotgun library construction
JOURNAL Anal. Biochem. 236 (1), 107-113 (1996)
MEDLINE 96207227
PUBMED 8619474
REFERENCE 2 (bases 1 to 1369)
AUTHORS Yu, W., Andersson, B., Worley, K.C., Muzny, D.M., Ding, Y., Liu, W.,
Ricafrente, J.Y., Wentland, M.A., Lennon, G., and Gibbs, R.A.
TITLE Large-scale concatenation cDNA sequencing
JOURNAL Genome Res. 7 (4), 353-358 (1997)
MEDLINE 97264341
PUBMED 9110174
REFERENCE 3 (bases 1 to 1369)
AUTHORS Yu, W., Sarginson, J., and Gibbs, R.A.
TITLE Direct Submission
JOURNAL Submitted (20-NOV-1997) Molecular and Human Genetics, Baylor
College of Medicine, One Baylor Plaza S930, Houston, TX 77030, USA
FEATURES
source 1.1369
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="23931"
/sex="female"
/tissue_type="brain"
/clone_lib="1NIB"
/dev_stage="infant"
/note="This clone is similar to Drosophila M. Abd-B gene
encoded by GenBank Accession Number L07835. The I.M.A.G.E.
consortium clone ID number is 23931"
<1.657
/note="similar to S-adenosylhomocysteine hydrolase"
/codon_start=1
/protein_id="AAB8189.1"
/db_xref="GI:2661083"
/translation="PRAMCGYGEVKGCCALKAALGAIVYITEIDPICALQACMDGFR
VKLINEVIRQDVIVITCTGNKNTVREHIDRMKNSCIVCNMGHNSITEIDVSLRPEL
TWERYSQVDHVIWPDGKRVLLAEGRLNLSCTVPTFVLSITATTOALALIELYNA
PEGRYKQDYLLPKMDEVASLHPSFDAHLTELTDQAKYILGNKNGPKPNYIRY"
```

CDS

```
BASE COUNT 370 a 314 c 286 g 399 t
ORIGIN
```

```
Query Match 49.7%; Score 644; DB 9; Length 1369;
Best Local Similarity 100.0%; Pred. No. 1.2e-189;
Matches 644; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 653 TGTGTGCTATGCTGAGCTAGGCAAGGCTGCTGCTCTCAAGCTCTTGAGACAA 712
Db 11 TGTGTGCTATGCTGAGCTAGGCAAGGCTGCTGCTCTCAAGCTCTTGAGACAA 70
```

QY	713	TTGCTACATTACCGAAATCGACCCCATCTGTGCTCTGCAGAGGCCCTGCATGGATGGGTCA	772
Db	71	TTGCTACATTACCGAAATCGACCCCATCTGTGCTCTGCAGAGGCCCTGCATGGATGGGTCA	130
QY	773	GGGTGTAAGCTTAATGAAGTCATCCGGCAAGTCGATGTCGTAATACTTGACAGGAA	832
Db	131	GGGTGTAAGCTTAATGAAGTCATCCGGCAAGTCGATGTCGTAATACTTGACAGGAA	190
QY	833	ATAAGATGTAGTGACACGGGAGCACCTTGGATCGCATGAATAACAGTGTATCGTATGCA	892
Db	191	ATAAGATGTAGTGACACGGGAGCACCTTGGATCGCATGAATAACAGTGTATCGTATGCA	250
QY	893	ATATGGGCCACTCCAACACAGAAATCGATGTGACCAAGCCTCCGCACCTCCGGAGCTGAGT	952
Db	251	ATATGGGCCACTCCAACACAGAAATCGATGTGACCAAGCCTCCGCACCTCCGGAGCTGAGT	310
QY	953	GGGAGCGAGTACGTTCTCAGGTGAGTGACATGTCACTGCGCAGATGGCAACGAGTTGTC	1012
Db	311	GGGAGCGAGTACGTTCTCAGGTGAGTGACATGTCACTGCGCAGATGGCAACGAGTTGTC	370
QY	1013	TCCTGGCAGAGGGTCTGCTACTCAATTTGAGCTGCTCCACAGTTCCCACTTTGTTCTGT	1072
Db	371	TCCTGGCAGAGGGTCTGCTACTCAATTTGAGCTGCTCCACAGTTCCCACTTTGTTCTGT	430
QY	1073	CCATCACAGCCACAACACAGGCTTTGGCACTGATAGACTCTATAATGCACCCGAGGGGC	1132
Db	431	CCATCACAGCCACAACACAGGCTTTGGCACTGATAGAACTCTATATGCACCCGAGGGGC	490
QY	1133	GATACAAGCAGGATGTGTACTTGCTTCTTAAGAAAAATGAGTAATACGTTGCCAGCTTGC	1192
Db	491	GATACAAGCAGGATGTGTACTTGCTTCTTAAGAAAAATGAGTAATACGTTGCCAGCTTGC	550
QY	1193	ATCTGCCATCATTTGATGCGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG	1252
Db	551	ATCTGCCATCATTTGATGCGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG	610
QY	1253	GACTCAACAAAAATGGGCCATTCAACCTAATTATTACAGATAC	1296
Db	611	GACTCAACAAAAATGGGCCATTCAACCTAATTATTACAGATAC	654

RESULT 14	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	
AY113501	AY113501	Drosophila melanogaster	RE58316	full insert	CDNA.	FLI_CDNA.	Drosophila melanogaster.	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Carlson, J., Champe, M., Chavez, C., Dorsett, V., Dresnek, D., Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Krommiller, B., Li, P., Liao, G., Miranda, A., Mungai, C. J., Nunoo, J., Pacleb, J., Paragas, V., Park, S., Patel, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Rubin, G. M. and Celniker, S.	Direct Submission	Submitted (16-MAY-2002)	Berkeley Drosophila Genome Project, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA	Sequence submitted by:

This clone was sequenced as part of a high-throughput process to sequence clones from *Drosophila* Gene Collection 1 (Rubin et al., 2000). The sequence has been subjected to integrity checks for sequence accuracy, presence of a polyA tail and contiguity within 100 kb in the genome. Thus we believe the sequence to

reflect accurately this particular cDNA clone. However, there are artifacts associated with the generation of cDNA clones that may have not been detected in our initial analyses such as internal priming, priming from contaminating genomic DNA, retained introns due to reverse transcription of unspliced precursor RNAs, and reverse transcriptase errors that result in single base changes. For further information about this sequence, including its location and relationship to other sequences, please visit our Web site (<http://fruitfly.berkeley.edu>) or send email to cdna@fruitfly.berkeley.edu.

Location/Qualifiers

```

source
1. .1792
/organism="Drosophila melanogaster"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
/map="89E10-89E10"
1. .1792
/gene="Ahcy89E"
/note="alignment with genomic scaffold AE003715"
/db_xref="FLYBASE:FBgn0015011"
221. .1699

```

```

CDS
221..1699
/gene="Ancy89E"
/note="Longest ORF"
/codon_start=1
/product="RE58316p"
/protein_id="AAM29506.1"
/db_xref="GI:21064553"
/db_xref="PIBASE:FBgn0015011"
/translation="MAKMPETFEADLSLADKTAVKKSSIEARFSDVSNCSFSCTCT
GSSDEEDVSPKDNHQNSAGGTDFCVKSISKSAGFREIEIAESEMPGIMTLRKRAKD
EKPLKGANIYGCETHVNAQSAVLIEITVLQCATVRMAACNIYSTQNAVAAALAEAGIRI
FAMRGETEETEEFWWCILDRAIYSDGWPNLIDDGADATHLMKKYPDYFAIRGIVEES
VTGVHRLVMSKSGCKLTPAIVNDSVTKKPEDFYTCRDSILDSLKRITDIMFGGKQ
VVICGYDVGKGAQSLKGGQCIYVTEYDPICALQAAADGFRVRLNERTVDVVV
TATGNKNVITRPHMNRMKNGCILCMGHSCSELDVNGLHTPELTWERVRQVDHIRWP
DGRMIIILAEGRVNLSCSTISSFVSVASSQALALIELFSAPGRYKSDVYLLPKKM
DEVVASLHLATFDALHTELTDQSKFMGLNKAQPFKANYRRY"
BASE COUNT 463 a 425 c 483 g 421 t
ORIGIN

```

Query Match	48.7%;	Score 631;	DB 3;	Length 1792;
Best Local Similarity	68.5%;	Pred. No. 1.4e-185;		
Matches 887; Conservative	0;	Mismatches 405;	Indels 3;	Gaps 1;

QY	1	TCCAAGGGCAGCAGCAATTTCTGTGAAGAACATCAAGCAGGCGAATTTGCAGCCGG	60
Dd	404	TCCGCCGAGGACACTGACTTTTGGGTGAAGACATCTCGAAGAGTGCAATTGGAAAGCCG	463
QY	61	GAGATTGAGATTGCAGAGCAAGACATGTCGTCTGTATTTCATCAGGAACGTGCTCAG	120
Dd	464	GAGATCGAGATCGCCGAGTCGAGATGCCGGCATCATGACTCTGAGGAAAAGAGCGAAG	523
QY	121	GGGGAAGACCCCTTGCGTGTGCTAAATAAGTGGGCTGTACACACATCACAGCCAGACA	180
Dd	524	GATGAGAAGCCCCCTAAAGGGTGCCAAATATCTCGGATGCACCACCGTCAATGCTCAGTCG	583
QY	181	GCGGTGTGATTGAGACACTCTGTGCCCTGGGGGCTCAGTGCCGCTGTCTGTTGTAAC	240
Dd	584	GCAGTGTGATCGAGACCCCTGTCCAACTGGCGCCACAGTTGCTGGGCTGCCGCAAC	643
QY	241	ATCTACTCACTCAGAATGAAGTAGCTGCAGCACTGGCTGAGGCTGAGTTGCAGTGTTTC	300
Dd	644	ATTATTTCCACACAAAAACGAGTTGCCCGCTCTGTGGCAGAGCGGGGAATTCGATCTTC	703
QY	301	GCTTGGAGGGCGAGTCAGAAAGATGACTTCTGTTGTTGATTGACCGCTGTGTGAACATG	360
Dd	704	GCTTGGCGCGAGAGACGAGAGAGAGTTCTGTGTGCTTGGACAGGGCCATCTACTCC	763
QY	361	GATGGTGGCAGGCCAACATGATCCTGATGATGGGGGAGAGACTTAACCCACTGGGTTAT	420
Dd	764	GACGGCTGGCAGCCGAACCTAATCTTGACGACGCGCGGCGATGCCACGCACCTATGCTC	823
QY	421	AAGAAGTATCCAACAGCTGTTAAGAAGATCCAGGCAATTGTGGAAGAGAGCGTACTGCT	480

||||| 824 AAGAGTACCCGACTACTTCAAGGCCATTCGGGGCATCGTGAGAGAAAGTGTGACCGGG 883
QY 481 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTGTCCGGCCATGAACGTC 540
Db 884 GTGCACCGGCTGTATCATGCTGTCAAGGGCGGAAACTTACTGTCCGGCCATCAACGTT 943
QY 541 AATGATCTGTATACCAACAAGAGTTGATTAATCTGTACTGTCTGCCGAGATCCATTG 600
Db 944 AACGACTCAGTACAGCAAGCAAGTTGATTAATCTGTACTGTCTGCCGAGATCCATTG 1003
QY 601 GATGCGCTGAAGAGACCAAGATGTGATGTTGGTGGGAAACAAGTGTGTGTGTGTC 660
Db 1004 GACAGTCTGAAGACCAAGATGTGATGTTGGTGGGAAACAAGTGTGTGTGTGTC 1063
QY 661 TATGTGAGTGAAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
Db 1064 TACGGT 1123
QY 721 ATTACGGAATCGACCCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
Db 1124 GTTACGGAATCGACCCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1183
QY 781 AAGCTAATGAAGTCAATCCGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
Db 1184 CGGCTCAAGAGT 1243
QY 841 GTAGTACACGGGACCACTGT 900
Db 1244 GTTATTAACAGGATCAATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1303
QY 901 CACTCCAACACAGAAATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
Db 1304 CATTCCTGCTCGGAGATGT 1363
QY 961 GTACGTTCTCAGGTGACCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
Db 1364 GTCCGTTCTCAGGTGACCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1423
QY 1021 GAGGTCGTCTACTCAATTTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
Db 1424 GAGGTCGTCTACTCAATTTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1483
QY 1081 GCCACAACACAGGCTTTGGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140
Db 1484 TCATCCACCCAGGCTTTGGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1540
QY 1141 CAGGATGT 1200
Db 1541 TCGATGT 1600
QY 1201 TCATTGATGCGCCACTTACAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260
Db 1601 ACCTTGATGCGCCACTTACAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1660
QY 1261 AAAAATGGGCAATCAACCTAATTAATTAACAGATA 1295
Db 1661 AAGGCCGGGCTTTTAAAGCAATTAATTAACAGATA 1695
RESULT 15
LOCUS BC003631 1323 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, similar to S-adenosylhomocysteine hydrolase-like 1,
clone IMAGE:3010755, mRNA, partial cds.
ACCESSION BC003631
VERSION BC003631.1 GI:13177702
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1323)

AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (26-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubln Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
Contact: amadan@systemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
Greene, Mark Kettelman and Anuradha Madan
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 3 Row: b Column: 6.
Location/Qualifiers
1. 1323
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3010755"
/tissue_type="Muscle, rhabdomyosarcoma"
/clone_id="NIH_MGC_17"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
<1. 611
/codon_start=3
/product="Similar to S-adenosylhomocysteine hydrolase-like
1"
/protein_id="AAH03631.1"
/db_xref="GI:13177703"
/translation="HEALGAIYITEIDPICALQACMDGFRVYKLEIVRODVIT
CTGNKIVYTRHLDRLMKNSCIYCNHNSNTEIDVTSRTPETLWVRVRSQVDHWPD
GKRVLLAEGRLLNSCSTVPPVLSITATQALALIELVNAPEGRYKODVYLLPKKM
DEVYASLHLPSPDHLTELTDQAKYLGINKNGPKPNYRY"
BASE COUNT 365 a 303 c 271 g 384 t
ORIGIN
Query Match 46.2%; Score 599; DB 9; Length 1323;
Best local Similarity 100.0%; Pred. No. 1.4e-175;
Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 698 AAGCTCTTGAGCAATGTCTACATTACCGAAATCGACCCCATCTGTGTGTGTGTGTGTGT 757
Db 10 AAGCTCTTGAGCAATGTCTACATTACCGAAATCGACCCCATCTGTGTGTGTGTGTGTGT 69
QY 758 GCATGATGGTTCAGGGTGTAAAGCTAATGAATCATCCGGCAAGTGTGTGTGTGTGTGT 817
Db 70 GCATGATGGTTCAGGGTGTAAAGCTAATGAATCATCCGGCAAGTGTGTGTGTGTGTGT 129
QY 818 TAACCTGCACAGAAATAAGATGTAGTACACGGGAGACACTTGGATCGCATGAAAAACA 877
Db 130 TAACCTGCACAGAAATAAGATGTAGTACACGGGAGACACTTGGATCGCATGAAAAACA 189
QY 878 GTGTATCGTATGCAATATGGGCCACTCCACACAGAAATCGATGTGACAGCCTCCGCA 937
Db 190 GTGTATCGTATGCAATATGGGCCACTCCACACAGAAATCGATGTGACAGCCTCCGCA 249
QY 938 CTCCGAGCTGAGCTGGAGGAGTACGTTCTCAGGTGAGGACCATGTGTGTGTGTGTGTGT 997
Db 250 CTCCGAGCTGAGCTGGAGGAGTACGTTCTCAGGTGAGGACCATGTGTGTGTGTGTGTGT 309
QY 998 GCAAACGAGTTCCTCCTGTGACAGAGGCTCTACTCAATTTGAGTGTCTCCACAGTTC 1057
Db 310 GCAAACGAGTTCCTCCTGTGACAGAGGCTCTACTCAATTTGAGTGTCTCCACAGTTC 369
QY 1058 CCACCTTTGTCTGTCCATCACAGCCACACACAGGCTTTGGCACCTGTAGAACTCTATA 1117

Db 370 CCACCTTGTCTGTCCATCACAGCCACACAGGCTTTGGCACTGATAGAACTCTATA 429
QY 1118 ATGCACCCGAGGGCGATACAAGCAGATGTGTACTTGTCTCTAAGAAAATGATGAAT 1177
|||||
Db 430 ATGCACCCGAGGGCGATACAAGCAGATGTGTACTTGTCTCTAAGAAAATGATGAAT 489
QY 1178 ACGTGGCAGCTGTGCACTCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACC 1237
|||||
Db 490 ACGTGGCAGCTGTGCACTCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACC 549
QY 1238 AAGCAAAATATCTGGGACTCAACAAAATGGGCCATTCAAAACCTAATTATTACAGATAC 1296
|||||
Db 550 AAGCAAAATATCTGGGACTCAACAAAATGGGCCATTCAAAACCTAATTATTACAGATAC 608

Search completed: April 21, 2003, 02:39:26
Job time : 3367.48 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2003, 21:15:39 ; Search time 275.45 Seconds
(without alignments)
10595.712 Million cell updates/sec

Title: US-09-782-051-1_COPY_549_1844

Perfect score: 1296
Sequence: 1 tccaagcgagcagcaattt.....aacctaattatcacagatac 1296

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 segs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*

1:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1296	100.0	2563	19	AAV28617
2	1296	100.0	3634	23	ABV23195
3	1296	100.0	3634	23	ABV29032
4	1294.4	99.9	3616	20	AAV73924
5	815.8	62.9	5030	23	AAV76216
6	676.6	52.2	1566	23	ABL03693
7	628.4	48.5	1776	23	ABL29787
8	486.2	37.5	4406	23	ABL03692
9	461.4	35.6	553	24	ABL79045

C	10	422	32.6	504	23	ABV44141	Human prostate exp
C	11	406.6	31.4	5537	23	ABL29786	Drosophila melanog
C	12	406.4	31.4	636	21	AAZ80766	Human colon cancer
C	13	402.4	31.0	2200	21	AAZ77823	Human cancer assoc
C	14	402.4	31.0	2658	23	ABV23126	Human prostate exp
C	15	402.4	31.0	2658	23	ABV28967	Human prostate exp
C	16	378.2	29.2	1592	23	ABL28157	Drosophila melanog
C	17	363.8	28.1	435	23	ABV14215	Human prostate exp
C	18	337	26.0	384	24	ABN93848	Gene #346 used to
C	19	314.4	24.3	2211	22	AAZ91228	Human S-adenosylho
C	20	313.2	24.2	4380	23	ABL28156	Drosophila melanog
C	21	245	18.9	939	23	AAV76217	DNA encoding novel
C	22	244.8	18.9	312	23	ABV35309	Human prostate exp
C	23	242.8	18.7	1727	21	AAZ46315	Human prostate exp
C	24	242.6	18.7	278	24	ABN73754	Arabidopsis thalia
C	25	241.6	18.6	1461	22	AAZ09278	Bovine embryonic g
C	26	236.2	18.2	1488	21	AAA39567	Trichomonas vagina
C	27	236.2	18.2	4411529	22	AAI99682	M. tuberculosis an
C	28	233.6	18.0	1461	22	AAZ09277	Trichomonas vagina
C	29	231.8	17.9	1728	21	AAZ39836	Arabidopsis thalia
C	30	228.8	17.7	1865	21	AAZ44037	Zea mays DNA fragm
C	31	226.4	17.5	1494	21	AAZ46817	Arabidopsis thalia
C	32	225.6	17.4	1422	22	AAZ65801	C glutamicum codin
C	33	225.6	17.4	1557	22	AAZ71862	Corynebacterium gl
C	34	225.6	17.4	1557	23	AAZ96121	C. glutamicum gene
C	35	225.6	17.4	1939	24	AAZ52394	Sahn gene encoding
C	36	225.6	17.4	349980	22	AAZ68526	C glutamicum codin
C	37	221	17.1	1812	17	AAZ29194	S-adenosylhomocyst
C	38	217.8	16.8	1812	13	AAZ029419	Nicotiana tabacum
C	39	214.6	16.6	1767	17	AAZ44513	Asparagus S-adenos
C	40	205.8	15.9	58909	22	AAZ28543	Genomic fragment #
C	41	202	15.6	656	21	AAZ14585	Aspergillus Oryzae
C	42	190	14.7	1877	18	AAV06354	Al-2-long (AL-21)
C	43	189.2	14.6	1941	17	AAZ44515	Arabidopsis S-aden
C	44	186	14.4	1396	22	AAZ71864	Corynebacterium gl
C	45	186	14.4	1396	23	AAZ95123	C. glutamicum gene

ALIGNMENTS

RESULT 1	AAV28617	standard; cDNA; 2563 BP.
ID	AAV28617	
AC	AAV28617;	
XX		
DT	21-AUG-1998	(first entry)
XX		
DE	Nucleotide sequence of the ODD4b5.3 enzyme.	
XX		
KW	DD4b9.3 enzyme; S-adenosyl-L-homocysteine hydrolase-type activity;	
KW	AHcy-type activity; stimulation; inhibition; dendritic cell; cancer;	
KW	autoimmune disease; transplantation; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	3..1847
FT		/*tag- a
FT		/product- "DD4b5.3 enzyme"
XX		
PN	W09814562-A1	
XX		
PD	09-APR-1998.	
XX		
PF	06-OCT-1997;	97WO-NZ00133.
XX		
PR	04-OCT-1996;	96NZ-0299507.
XX		
PA	(HART/) HART D N J.	
XX		
PI	Hart DNU;	

XX WPI: 1998-240074/21.
DR P-PSDB; AAW56097.
XX
PT New isolated S-adenosyl-L-homocysteine hydrolase enzyme - is used to
PT develop products which can be used in the treatment of e.g.
PT auto-immune disease, transplantations or cancers
XX
PS Claim 7; Fig 1; 33pp; English.
XX
CC This is the nucleotide sequence of the DD4b9.3 enzyme which has
CC S-adenosyl-L-homocysteine hydrolase (AHCY)-type activity. Its
CC products can be used to identify substances which have a stimulatory
CC or inhibitory effect on the enzyme activity. Such substances can be
CC used to modulate dendritic cell (DC) function and for immunomodulation.
CC They can be used in the treatment of e.g. autoimmune diseases,
CC transplantations or cancers. The products can also be used for
CC detection and diagnosis.
XX
SO Sequence: 2563 BP; 646 A; 604 C; 677 G; 636 T; 0 other;
Query Match 100.0%; Score 1296; DB 19; Length 2563;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCCAGGGGCGAGCAATTTCTGTGTGAAGAACATCAAGCAGGAGAAATTTGGAGCCGG 60
DB 549 TCCAGGGGCGAGCAATTTCTGTGTGAAGAACATCAAGCAGGAGAAATTTGGAGCCGG 608
QY 61 GAGATTGAGATTGCAGAGCAAGACATGTCTGTCTGATTTCACAGGAACGTGCTCAG 120
DB 609 GAGATTGAGATTGCAGAGCAAGACATGTCTGTCTGATTTCACAGGAACGTGCTCAG 668
QY 121 GGGAGAGAGCCCTTGCTGTGTGAAGAACATCAAGCAGGAGAAATTTGGAGCCGG 180
DB 669 GGGAGAGAGCCCTTGCTGTGTGAAGAACATCAAGCAGGAGAAATTTGGAGCCGG 728
QY 181 GCGGTGTGATTGAGACACTCTGTGCCCCGGGGCTCAGTGGCGCTGCTGCTGTATAC 240
DB 729 GCGGTGTGATTGAGACACTCTGTGCCCCGGGGCTCAGTGGCGCTGCTGCTGTATAC 788
QY 241 ATCTACTCACTCAGAAATGAAGTACTGCAGCAGCTGGCTGAGGCTGGAGTTGCACTGTT 300
DB 789 ATCTACTCACTCAGAAATGAAGTACTGCAGCAGCTGGCTGAGGCTGGAGTTGCACTGTT 848
QY 301 GCTTGGAGGGCGAGTCAAGAGTACTTCTGTGTGATTGACCCGCTGTGTGAACATG 360
DB 849 GCTTGGAGGGCGAGTCAAGAGTACTTCTGTGTGATTGACCCGCTGTGTGAACATG 908
QY 361 GATGGGTGGCAGGCCAATGATCTGATGATGGGGAGAGCTTAACCCACTGGTTAT 420
DB 909 GATGGGTGGCAGGCCAATGATCTGATGATGGGGAGAGCTTAACCCACTGGTTAT 968
QY 421 AAGAGTATCCCAACGTGTTAAGAGATCCGAGGCAATTTGAGAGAGAGCGTGACTGGT 480
DB 969 AAGAGTATCCCAACGTGTTAAGAGATCCGAGGCAATTTGAGAGAGAGCGTGACTGGT 1028
QY 481 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTGTTCCGGCCATGAACGTC 540
DB 1029 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTGTTCCGGCCATGAACGTC 1088
QY 541 AATGATTCTGTATCCAAACAGAGTTGATTAATCTGTACTGCTCCGAGAAATCCATTG 600
DB 1089 AATGATTCTGTATCCAAACAGAGTTGATTAATCTGTACTGCTCCGAGAAATCCATTG 1148
QY 601 GATGGCTGAAGAGGACACAGATGATGTTTGGTGGAAACAAGTGTGTGTGGC 660
DB 1149 GATGGCTGAAGAGGACACAGATGATGTTTGGTGGAAACAAGTGTGTGTGGC 1208
QY 661 TATGTTGAGTAGGCAAGGGCTGTGTGCTCTCAAAAGCTCTTGAGAGCAATGTCTAC 720
DB 1209 TATGTTGAGTAGGCAAGGGCTGTGTGCTCTCAAAAGCTCTTGAGAGCAATGTCTAC 1268

QY 721 ATTAACGGAATCGACCCCATCTGTGCTCTGCAGGCCCTGCATGGATGGGTTCAAGGTTGTA 780
DB 1269 ATTAACGGAATCGACCCCATCTGTGCTCTGCAGGCCCTGCATGGATGGGTTCAAGGTTGTA 1328
QY 781 AAGCTAAATGAAGTCAATCCGGCAAGTGCATGTCGTAATTAATTCACAGAGAAATAGAT 840
DB 1329 AAGCTAAATGAAGTCAATCCGGCAAGTGCATGTCGTAATTAATTCACAGAGAAATAGAT 1388
QY 841 GTAGTGACACGGGAGCAGCTTGATCCGATGAAACAGAGTTGATCCGATGCAATATGGGC 900
DB 1389 GTAGTGACACGGGAGCAGCTTGATCCGATGAAACAGAGTTGATCCGATGCAATATGGGC 1448
QY 901 CACTCAACACAGAAATGATGTGACCAAGCCTCCGCACTCCGAGAGCTGACGTGGAGCGA 960
DB 1449 CACTCAACACAGAAATGATGTGACCAAGCCTCCGCACTCCGAGAGCTGACGTGGAGCGA 1508
QY 961 GTACGTTCTCAGGTGACCATGTATCTGCGCAGATGGCAAGAGTTGCTCCTGCGCA 1020
DB 1509 GTACGTTCTCAGGTGACCATGTATCTGCGCAGATGGCAAGAGTTGCTCCTGCGCA 1568
QY 1021 GAGGTCGCTCACTCAATTTGAGCTGCTCCACAGTTCCACCTTGTCTGTCCATCACA 1080
DB 1569 GAGGTCGCTCACTCAATTTGAGCTGCTCCACAGTTCCACCTTGTCTGTCCATCACA 1628
QY 1081 GCCACACACAGGCTTTGGCAGCTGATAGAACTCTATAATGACCCGAGGGCGATACAG 1140
DB 1629 GCCACACACAGGCTTTGGCAGCTGATAGAACTCTATAATGACCCGAGGGCGATACAG 1688
QY 1141 CAGGATGTGTAATCTGCTTCCCTAAGAAATGATGATGATGATGATGATGATGATGAT 1200
DB 1689 CAGGATGTGTAATCTGCTTCCCTAAGAAATGATGATGATGATGATGATGATGATGAT 1748
QY 1201 TCATTTGATGCCCACTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGAGCTCAAC 1260
DB 1749 TCATTTGATGCCCACTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGAGCTCAAC 1808
QY 1261 AAAAATGGCCATTCAAACTTAATTTACAGATAC 1296
DB 1809 AAAAATGGCCATTCAAACTTAATTTACAGATAC 1844
RESULT 2
ABV23195
ID ABV23195 standard; cDNA; 3634 BP.
XX AC ABV23195;
XX AC
XX DT 16-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 23186.
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN WO200160860-A2.
XX PD 23-AUG-2001.
XX PE 20-FEB-2001; 2001WO-US05171.
XX PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege WO, Monahan JE.
XX

DR WPI: 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer

XX Claim 1; Page 4168-4169; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX Sequence 3634 BP; 876 A; 966 C; 869 G; 918 T; 5 other;

Query Match 100.0%; Score 1296; DB 23; Length 3634;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TCCAAGGCGACGACCAATTTCTGTGTGAAGAACATCAAGCAGGAGATTTGAGCGCCG 60
DB 619 TCCAAGGCGACGACCAATTTCTGTGTGAAGAACATCAAGCAGGAGATTTGAGCGCCG 678
OY 61 GAGATTGAGATTGCGACGACGACATGTCTGCTGTGATTTCACTCAGGAAACGCTCTAG 120
DB 679 GAGATTGAGATTGCGACGACGACATGTCTGCTGTGATTTCACTCAGGAAACGCTCTAG 738
OY 121 GGGGAGAGCCCTTGGCTGCTCTAAATAGTGGGCTGTACACACATCACAGCCGACGACA 180
DB 739 GGGGAGAGCCCTTGGCTGCTCTAAATAGTGGGCTGTACACACATCACAGCCGACGACA 798
OY 181 GCGGTGTGATTGAGACACTCTGTGCCCTGGGGCTCAGTGCCTGTCTGTGTAAC 240
DB 799 GCGGTGTGATTGAGACACTCTGTGCCCTGGGGCTCAGTGCCTGTCTGTGTAAC 858
OY 241 ATCTACTCACTCAGATGAGTAGCTGCAGCAGTGGCTGAGGCTGGAGTGCAGTGTTC 300
DB 859 ATCTACTCACTCAGATGAGTAGCTGCAGCAGTGGCTGAGGCTGGAGTGCAGTGTTC 918
OY 301 GCTTGAAGGGGCGAGTCAAGATGACTTCTGTGTGTGATTTGACCGCTGTGTAACATG 360
DB 919 GCTTGAAGGGGCGAGTCAAGATGACTTCTGTGTGTGATTTGACCGCTGTGTAACATG 978
OY 361 GATGGGTGCGAGGCGCAATGATCTGTGATGAGGAGGAGACTTAACCCACTGGGTTTAT 420
DB 979 GATGGGTGCGAGGCGCAATGATCTGTGATGAGGAGGAGACTTAACCCACTGGGTTTAT 1038
OY 421 AAGAAGTATCCAAACGTTTAAAGAGATCCGAGCATTTGGAAGAGAGCGTGAATGCT 480
DB 1039 AAGAAGTATCCAAACGTTTAAAGAGATCCGAGCATTTGGAAGAGAGCGTGAATGCT 1098
OY 481 GTTCACAGGCTGTATCAGTCTCCAAAGCTGGGAGAGCTGTGTCCGGCCATGAACGTC 540
DB 1099 GTTCACAGGCTGTATCAGTCTCCAAAGCTGGGAGAGCTGTGTCCGGCCATGAACGTC 1158
OY 541 AATGATTCTGTACCAACAGAGTTTGATTAACCTGTACTGCTCCGAGAAATCCATTTTG 600
DB 1159 AATGATTCTGTACCAACAGAGTTTGATTAACCTGTACTGCTCCGAGAAATCCATTTTG 1218
OY 601 GATGGCTGAAGAGAGACAGATGATGTTTGTGGGAAACAAGTGTGTGTGTCG 660
DB 1219 GATGGCTGAAGAGAGACAGATGATGTTTGTGGGAAACAAGTGTGTGTGTCG 1278

OY 661 TATGTTGAGGTAGGCAAGGGCTGTGTGCTGTCTCAAGCTCTTGAGCAATGTCTAC 720
DB 1279 TATGTTGAGGTAGGCAAGGGCTGTGTGCTGTCTCTCAAGCTCTTGAGCAATGTCTAC 1338
OY 721 ATTACCGAAATCGACCCCATCTGTGCTCTGCAAGGCTGATGGGTTGAGGGTGTGTA 780
DB 1339 ATTACCGAAATCGACCCCATCTGTGCTCTGCAAGGCTGATGGGTTGAGGGTGTGTA 1398
OY 781 AAGCTAATGAGTCAATCCGCAAGTGTGATGTCTTAATCTGCACAGGAAATAGAT 840
DB 1399 AAGCTAATGAGTCAATCCGCAAGTGTGATGTCTTAATCTGCACAGGAAATAGAT 1458
OY 841 GTAGTACACGCGGACGACCTTGGATCGCATGAAAAACAGTTGTATGCAATATGGC 900
DB 1459 GTAGTACACGCGGACGACCTTGGATCGCATGAAAAACAGTTGTATGCAATATGGC 1518
OY 901 CACTCCAACACAGAAATCGATGACACAGCCTCCGACCTCCGAGCTGACGTGGAGCGA 960
DB 1519 CACTCCAACACAGAAATCGATGACACAGCCTCCGACCTCCGAGCTGACGTGGAGCGA 1578
OY 961 GTAGTCTCAGGTGAGGACCATCTGTGCGCAGATGGCAACGATTTGCTCTGCGCA 1020
DB 1579 GTAGTCTCAGGTGAGGACCATCTGTGCGCAGATGGCAACGATTTGCTCTGCGCA 1638
OY 1021 GAGGTGCTTACTCAATTTGACCTGCTCCACAGTCCCACTTTGTTCTGTCATCACA 1080
DB 1639 GAGGTGCTTACTCAATTTGACCTGCTCCACAGTCCCACTTTGTTCTGTCATCACA 1698
OY 1081 GCCACAACACAGGCTTTGGCAGTGTAGAACTCTATATGCAACCCGAGGGGATACAG 1140
DB 1699 GCCACAACACAGGCTTTGGCAGTGTAGAACTCTATATGCAACCCGAGGGGATACAG 1758
OY 1141 CAGGATGTGACTTGTCTCTTGAAGAAATGATGATAGTGTGCGCAGCTTGACATGCGA 1200
DB 1759 CAGGATGTGACTTGTCTCTTGAAGAAATGATGATAGTGTGCGCAGCTTGACATGCGA 1818
OY 1201 TCATTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGACATCAAC 1260
DB 1819 TCATTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGACATCAAC 1878
OY 1261 AAAAATGGGCCATTCAAACTTAATTTACAGATAC 1296
DB 1879 AAAAATGGGCCATTCAAACTTAATTTACAGATAC 1914

RESULT 3
ABV29032
ID ABV29032 standard; cDNA; 3634 BP.
XX
AC ABV29032;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 29023.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX Schlegel R, Endege WO, Monahan JE;
 XX WPI; 2001-662795/76.
 PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer
 XX
 PS Claim 1; Page 6147-6148; 11750bp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
 CC
 XX Sequence 3634 BP; 876 A; 966 C; 869 G; 918 T; 5 other;
 SQ
 Query Match 100.0%; Score 1296; DB 23; Length 3634;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1159 AATGATCTCTGTACCAACAGAGTTTGATACTGTGCTGCTGCCGAGATCCATTG 1218
 QY 601 GATGCCCTGAAGAGGACACAGATGTGATGTTGGGAAACAAGTGTGTGTC 660
 Db 1219 GATGCCCTGAAGAGGACACAGATGTGATGTTGGGAAACAAGTGTGTGTC 1278
 QY 661 TATGTTGAGGTAGGCAAGGGCTGTGCTGCTCTCAAAAGCTCTTGAGCAATGTCTAC 720
 Db 1279 TATGTTGAGGTAGGCAAGGGCTGTGCTGCTCTCAAAAGCTCTTGAGCAATGTCTAC 1338
 QY 721 ATACCGAATCGACCCCATCTGCTCTGACAGGCTGCAATGAGTGGTCAAGGTGTA 780
 Db 1339 ATACCGAATCGACCCCATCTGCTCTGACAGGCTGCAATGAGTGGTCAAGGTGTA 1398
 QY 781 AAGCTAATGAAGTCAATCCGGCAGTCGATGTCGTAATTAATTCACACAGAAATAGAAAT 840
 Db 1399 AAGCTAATGAAGTCAATCCGGCAGTCGATGTCGTAATTAATTCACACAGAAATAGAAAT 1458
 QY 841 GTAGTGACACGGGACGACCTTGATCGCATGAAAAACAGTTGTATGTAATATGAGGC 900
 Db 1459 GTAGTGACACGGGACGACCTTGATCGCATGAAAAACAGTTGTATGTAATATGAGGC 1518
 QY 901 CACTCCAAACAGAAATGATGATGACAGCCCTCCGACCTCCGAGAGTGCAGTGGAGCGA 960
 Db 1519 CACTCCAAACAGAAATGATGATGACAGCCCTCCGACCTCCGAGAGTGCAGTGGAGCGA 1578
 QY 961 GTAGTTCTCAGGTGACCATGTCATCTGGCCAGATGGCAACAGTTGTCTCTCTGGA 1020
 Db 1579 GTAGTTCTCAGGTGACCATGTCATCTGGCCAGATGGCAACAGTTGTCTCTCTGGA 1638
 QY 1021 GAGGTCTCTACTCAATTTGAGCTGCTCCACAGTTCACCTTTGTCTCTCATGACA 1080
 Db 1639 GAGGTCTCTACTCAATTTGAGCTGCTCCACAGTTCACCTTTGTCTCTCATGACA 1698
 QY 1081 GCCACAACACAGGCTTTGGCAGTATGAACTCTATAATGCAACCCGAGGGCGATACAG 1140
 Db 1699 GCCACAACACAGGCTTTGGCAGTATGAACTCTATAATGCAACCCGAGGGCGATACAG 1758
 QY 1141 CAGATGTGTACTTGTCTCTTAAGAAATGATGATGATGATGATGATGATGATGATGAT 1200
 Db 1759 CAGATGTGTACTTGTCTCTTAAGAAATGATGATGATGATGATGATGATGATGATGAT 1818
 QY 1201 TCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGAGCTCAC 1260
 Db 1819 TCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGAGCTCAC 1878
 QY 1261 AAAAATGGCCATTCAAACCTAATATATACAGATAC 1296
 Db 1879 AAAAATGGCCATTCAAACCTAATATATACAGATAC 1914

RESULT 4
 AAV73924
 ID AAV73924 standard; DNA; 3616 BP.
 XX
 AC AAV73924;
 XX
 DT 04-MAR-1999 (first entry)
 XX
 DE Human SAHH DNA #1.
 XX
 KW S-adenosyl-5-homocysteine hydrolase; SAHH; human; drug screening;
 KW treatment; infection; cancer; autoimmune disease; detection; diagnosis;
 KW gene mapping; antisense; therapy; antagonist; immunoassay; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 56..1558
 FT /*tag= a
 FT /product= "SAHH"
 XX
 PN US5854023-A.

XX 29-DEC-1998.
PD
XX 17-JUL-1997; 97US-0896005.
PF
XX 17-JUL-1997; 97US-0896005.
PR
XX (INCY-) INCYTE PHARM INC.
PA
XX Corley NC, Hillman JL, Lal P, Shah P;
XX WPI; 1999-094906/08.
DR P-PSDB; AAW90061.
XX
XX Nucleic acid encoding human S-adenosyl-5-homocysteine hydrolase -
PT for production of recombinant enzyme, useful for diagnosis,
PT treatment and prevention of cancers, infections and autoimmune
PT diseases
XX
XX Disclosure; Fig 1A-I; 40pp; English.
PS
XX This sequence encodes a human S-adenosyl-5-homocysteine hydrolase (SAHH).
CC The SAHH protein can be used to generate specific antibodies and in drug
CC screening to identify specific binding agents. Antagonists of the
CC protein are used to treat or prevent a wide range of viral, bacterial,
CC fungal, parasitic, protozoal or helminthic infections, many cancers
CC (leukaemia, lymphoma or solid tumours), and many autoimmune diseases
CC (e.g. acquired immune deficiency syndrome, allergy, asthma, diabetes
CC mellitus, multiple sclerosis etc). All these conditions may be treated by
CC expressing antisense sequences, triplex-forming agents or ribozymes
CC directed against the nucleic acid. The nucleic acid and its fragments can
CC be used as probes or primers for detecting and quantifying gene
CC expression, for diagnosis or monitoring of disease, to identify genetic
CC variations, mutations or polymorphisms, in gene mapping and as antisense
CC therapeutics. Antibodies are used directly as antagonists, indirectly to
CC deliver active agents to SAHH-expressing cells, to diagnose and monitor
CC diseases in standard immunoassays, in competitive drug screens and to
CC isolate the protein from natural sources.
XX
XX Sequence 3616 BP; 1017 A; 782 C; 826 G; 991 T; 0 other;
SQ

Query Match 99.9%; Score 1294.4; DB 20; Length 3616;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1295; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCCAAGGCGCAGCAATTTCTGTGTGAGAACAATCAAGCAGGCGAGAAATTTGGACGCCGG 60
DB 260 TCCAAGGCGCAGCAATTTCTGTGTGAGAACAATCAAGCAGGCGAGAAATTTGGACGCCGG 319
QY 61 GAGATTGAGATTGACAGCAGCAAGACATGCTGCTGATTTCACTCAGGAAACGTGCTCAG 120
DB 320 GAGATTGAGATTGACAGCAGCAAGACATGCTGCTGATTTCACTCAGGAAACGTGCTCAG 379
QY 121 GGGGAGAGCCCTTGCTGTGCTAAATAAGTGGGCTGTACACACATCAAGCCAGACA 180
DB 380 GGGGAGAGCCCTTGCTGTGCTAAATAAGTGGGCTGTACACACATCAAGCCAGACA 439
QY 181 GCGGTGTGATTGAGACACTCTGTGCTGGGGGCTCAGTCCGCTGCTGCTGTTAAC 240
DB 440 GCGGTGTGATTGAGACACTCTGTGCTGGGGGCTCAGTCCGCTGCTGCTGTTAAC 499
QY 241 ATGTACTCACTCAGAATGAGTAGCTGACGACACTGGCTGAGGCTGAGTTGCACTGTTTC 300
DB 500 ATGTACTCACTCAGAATGAGTAGCTGACGACACTGGCTGAGGCTGAGTTGCACTGTTTC 559
QY 301 GCTTGAAGGGCGAGTCAGAGATGACTTCTGCTGTGTATGACCGCTGTGTAACATG 360
DB 560 GCTTGAAGGGCGAGTCAGAGATGACTTCTGCTGTGTATGACCGCTGTGTAACATG 619
QY 361 GATGGGTGCGAGCCCAACATGATCCTGTGATGATGGGGAGACTTAACCCACTGGGTTTAT 420
DB 620 GATGGGTGCGAGCCCAACATGATCCTGTGATGATGGGGAGACTTAACCCACTGGGTTTGT 679

QY 421 AAGAGTATCCAAACGTGTTTAAGAGATCCGAGGCATGTGTGAAGAGAGCGTGACTGCT 480
DB 680 AAGAGTATCCAAACGTGTTTAAGAGATCCGAGGCATGTGTGAAGAGAGCGTGACTGCT 739
QY 481 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGGAAGCTCTGTGTTCGGCCATGAACGTC 540
DB 740 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGGAAGCTCTGTGTTCGGCCATGAACGTC 799
QY 541 AATGATTCGTATCCAAACAGAGTTGATTAATCTGTACTGTGCGGAGATCCATTTTG 600
DB 800 AATGATTCGTATCCAAACAGAGTTGATTAATCTGTACTGTGCGGAGATCCATTTTG 859
QY 601 GATGGCCTGAAGAGACCAACAGATGTGATGTTGGTGGGAAACAAGTGGTGTGTGGC 660
DB 860 GATGGCCTGAAGAGACCAACAGATGTGATGTTGGTGGGAAACAAGTGGTGTGTGGC 919
QY 661 TATGTTGAGGTAGGCAAGGGCTGCTGTGCTGCTCAAGCTCTTGAGCAATGCTTAC 720
DB 920 TATGTTGAGGTAGGCAAGGGCTGCTGTGCTGCTCAAGCTCTTGAGCAATGCTTAC 979
QY 721 ATTACCGAATCCAGCCCATCTGTGCTCTGACAGGCTGCATGATGGGTTACAGGCTGTA 780
DB 980 ATTACCGAATCCAGCCCATCTGTGCTCTGACAGGCTGCATGATGGGTTACAGGCTGTA 1039
QY 781 AAGCTAAATGAAGTCATCCGGCAAGTCGATGCTAATACTTGACAGGAAATAAGAT 840
DB 1040 AAGCTAAATGAAGTCATCCGGCAAGTCGATGCTAATACTTGACAGGAAATAAGAT 1099
QY 841 GTAGTGACACGGGACACTTGGATCGCATGAAACAGTTGTATGCTATGCAATATGGGC 900
DB 1100 GTAGTGACACGGGACACTTGGATCGCATGAAACAGTTGTATGCTATGCAATATGGGC 1159
QY 901 CACTCCAACACAGAAATCGATGTGACACAGCCCTCCGCACTCCGAGCTGAGCTGGAGCGA 960
DB 1160 CACTCCAACACAGAAATCGATGTGACACAGCCCTCCGCACTCCGAGCTGAGCTGGAGCGA 1219
QY 961 GTACGTTCTCAGGTGGACCATGTCTATCTGCGCCAGATGGCAACGAGTTGCTCTGCA 1020
DB 1220 GTACGTTCTCAGGTGGACCATGTCTATCTGCGCCAGATGGCAACGAGTTGCTCTGCA 1279
QY 1021 GAGGTGCTCTACTCAATTTGAGCTGCTCCACAGTCCACCTTGTGCTGTCATCACA 1080
DB 1280 GAGGTGCTCTACTCAATTTGAGCTGCTCCACAGTCCACCTTGTGCTGTCATCACA 1339
QY 1081 GCCACAACACAGGCTTTGGCACTGATAGAACTCTAATGACCCGAGGGCGATACAG 1140
DB 1340 GCCACAACACAGGCTTTGGCACTGATAGAACTCTAATGACCCGAGGGCGATACAG 1399
QY 1141 CAGGATGTCTACTTCTCTTAAGAAATGATGAATACGTGCGCAGCTTGCACTTGCCA 1200
DB 1400 CAGGATGTCTACTTCTCTTAAGAAATGATGAATACGTGCGCAGCTTGCACTTGCCA 1459
QY 1201 TCATTGTATGCCCACTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAAC 1260
DB 1460 TCATTGTATGCCCACTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAAC 1519
QY 1261 AAAAATGGCCATCAAACTAATATATACAGATAC 1296
DB 1520 AAAAATGGCCATCAAACTAATATATACAGATAC 1555

RESULT 5
AAS76216/c
ID AAS76216 standard; cDNA; 5030 BP.
XX AAS76216;
AC
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #12020.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.
OS
XX
PN W0200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Dimaanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
XX P-PSDB; ABG12029.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 1; SEQ ID No 12020; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics, as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SO Sequence 5030 BP; 1432 A; 1137 C; 1178 G; 1283 T; 0 other;

Query Match 62.9%; Score 815.8; DB 23; Length 5030;
Best Local Similarity 78.8%; Pred. No. 1.4e-254;
Matches 1023; Conservative 0; Mismatches 272; Indels 4; Gaps 4;

QY 1 TCCAGGCGCAGCAGCAATTTCTGTGTGAGAACATCAAGCAGCAATTGGACGCGG 60
DB 4469 TCTAAGGGAAGCAGTACTTGTGTGTAAGAACATCAAGCAGCAAGTTGGACGAGA 4410

QY 61 GAGATTGAGATTGAGAGACATGTCTGTGATTCATCTCAGGAACGTGCTCAG 120
DB 4409 GAAATGGAATGCTGAGACAGAAATGCTGATGATGCTTGAAGAGAGAGCTCA 4350

QY 121 GGGGAGAGAGCCCTTGGCTGTCTAAATAGTGGGCTGTACACATCAGACCCAGACA 180
DB 4349 GGAGAGAAAGCCTTGGCTGTGAGCAAAATCGTGGGTTGCACACATCAGCTCAGACT 4290

QY 181 GCGGTGTGATTGAGACATCTGTGCTGGGGCTCAGTGGCTGCTGCTGTGTAC 240
DB 4289 GCTGTCTTATGGAACCTGTGGTCTCTGGGGGCCAGTGGCGATGGGCTGCTGCAAC 4230

QY 241 ATCTACTCACTCAGAAATGAGTAGCTGAGCAGCTGGCTGAGGCTGAGTTGCACTGCTC 300
DB 4229 ATCTATTCACCTCAATGAAGTGGCTGCTGCTAGCAGAAAGTGGATTCTCTGTTT 4170

QY 301 GCTTGAAGGGGAGTCAGAGATGACTTCTGTGTGTATTGACCGCTGTGTGAACATG 360
DB 4169 GCCTGGAAGGAGAGTCAGAGATGACTTGTGTGTATTGACAGATGATGATGATG 4110

QY 361 GATGGTGGCAGGCAACATGATCTGTGATGATGGGGAGACTTAACCCACCTGGTTAT 420
DB 4109 GAGGGCTGGCAGGCAACATGATCTGTGATGATGAGGGGATCTTACCCACTGATTTAT 4050

QY 421 AAGAAGTATCCAAACGTGTTAAGAGATCCGAGCATTTGGAAGAGAGCGTACTGCT 480
DB 4049 AAAAAGTATCCCAACATGTTAAGAAATCAAGGCGATAGTAGAGAGAGTGTACTGGA 3990

QY 481 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTTCGGCCATGAACGTC 540
DB 3989 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTTCGGCCATGAACGTC 3930

QY 541 AATGATCTGTATCCAAACAGAGTTGATTAACCTGTACTGCTGCGGAGAAATCCATTG 600
DB 3929 AATGATCTGTATCCAAACAGAGTTGATTAACCTGTACTGCTGCGGAGAAATCCATTG 3870

QY 601 GATGGCTGAAGAGAGACACAGATGATGATGTTGTGGGGAACAAGTGTGTGTGCGC 660
DB 3869 GATGGCTGAAGAGAGACACAGATGATGATGTTGTGGGGAACAAGTGTGTGTGCGC 3810

QY 661 TATGCTGAGTAGGCAAGGCTGCTGTGCTGCTCTCAAGCTCTTGAGCAATGTCTAC 720
DB 3809 TATGAGAGGTGGGGAAGGCTGCTGTGCTGCTGCTGCAAGCCATGGGCTCATTTGTAT 3750

QY 721 ATTACGAATCGACCCCACTGTGCTGTGCTGCTGAGGCTGATGATGGTTCAAGGCTGA 780
DB 3749 GTAACTGAATGACCCCACTGTGCTGTGCTGCTGAGGCTGATGATGGTTCAAGGCTGA 3690

QY 781 AAGCTAATGAGTCACTGCGCAAG-TGATGTGCTAATTAATCTGCACAGAAATAGAA 839
DB 3689 AATTAATGAGTCACTGCGCAAG-TGATGTGCTAATTAATCTGCACAGAAATAGAA 3630

QY 840 TGTAGTGACACGAGGAGACCTTGATCGCATGAAACAGTTGATGATGATGATGATG 899
DB 3629 TGTAGTGACACGAGGAGACCTTGATCGCATGAAACAGTTGATGATGATGATGATG 3570

QY 900 CCACTCCACACAGAAATGATGTGACACAGGCTCGCACTCC-GGAGCTGACGTGGAGC 958
DB 3569 ACATTCACACAGAGATGATGTGACACAGGCTCGCACTCC-GGAGCTGACGTGGAGC 3510

QY 959 GAGTACGTTCTCAGGTGACATGTCATCTGCGCAGATGGCAACAGATGTCCTCTG 1018
DB 3509 GAGTACGTTCTCAGGTGACATGTCATCTGCGCAGATGGCAACAGATGTCCTCTG 3450

QY 1019 CAGAGGCTGCTACTCAATTTGAGCTGCTCCACAGTCCACCTTTGTTCTGTCATCA 1078
DB 3449 CAGAGGCTGCTACTCAATTTGAGCTGCTCCACAGTCCACCTTTGTTCTGTCATCA 3390

QY 1079 CAGCACAACAC-AGGCTTTGGCAGTATGAGTCACTCAATTTGAGTCCAGGCGATAC 1137
DB 3389 CTGCTACTACTCAAGGCTCTGCTGCTGATAGAGCTTACATGCTCTGAGGCTGCTAT 3330

QY 1138 AAGCAGATGTGCTGCTGCTTCAAGAAATGATGATGATGATGATGATGATGATGAT 1196
DB 3329 AAGCAGATGTGCTGCTGCTTCAAGAAATGATGATGATGATGATGATGATGATGAT 3270

QY 1197 GCCATCTTGTATGCTGCTGCTTCAAGAGCTGACAGATGACCAAGCAAAATATCTGGACT 1256
DB 3269 GCCATCTTGTATGCTGCTGCTTCAAGAGCTGACAGATGACCAAGCAAAATATCTGGACT 3210

QY 1257 CAACAAAATGGGCAATCAAACTAATTATTACAGATA 1295
DB 3209 CAATAGAATGGGCAATCAAACTAATTATTACAGATA 3171

RESULT 6
ABL03693
ID ABL03693 standard; cDNA; 1566 BP.

XX ABL03693;
AC
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 5561.
DE
XX Drosophila; developmental biology; cell signalling; insecticide;
KM pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
OS
XX WO200171042-A2.
PN
XX 27-SEP-2001.
PD
XX 23-MAR-2001; 2001WO-US09231.
PF
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
PA
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX WPI; 2001-656860/75.
DR P-PSDB; ABB59590.
DR
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions
XX
XX Claim 1; SEQ ID NO 5561; 21pp + Sequence Listing; English.
PS
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (AB57737-AB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
CC
XX
XX Sequence 1566 BP; 360 A; 438 C; 481 G; 287 T; 0 other;
SQ
Query Match 52.2%; Score 676.6; DB 23; Length 1566;
Best Local Similarity 70.6%; Pred. No. 1.7e-209;
Matches 917; Conservative 0; Mismatches 379; Indels 3; Gaps 1;

Db 565 TTGCGCTGGCGCGGCGAGACGAGAGAGACTTCTGTGTGATAGATCGCTGCTAAC 624
QY 358 ATGGATGGGTGGCAGGCCAATGATCTGTGATGATGGGGAGACTTAACCCACTGGGTT 417
Db 625 GCCGAGAACTGGCAGCCAAACATGATCCTGGACGATGGCGGATGCCACGCACTTGATG 684
QY 418 TATAGAAATATCCAAAGCTGTTAAGAGATCCGAGCATGTGGAAGAGAGCGTGAAT 477
Db 685 CTAAAAAGTACCCGACCATGTTCAGTTGGTCAAGGGAATCGTAGAGAGCGTCAAC 744
QY 478 GGTGTTACAGGCGCTGATGAGCTCTCCAAAGCTGGGAAGCTCTGTCTCCGGCCATGAC 537
Db 745 GGAGTGCACCGCTCTATACAGCTCTTAAGCCGCGCAAGCTGACAGTGCAGGATGAAT 804
QY 538 GTCAATGATCTGTTACCAACAGAGTTGATTAACCTGTACTGCTGCCGAGATTCATT 597
Db 805 GTCAAGATTCGGTGAACAAAGACCAAAATTTGACAAACCTTACAGCTGCAAGAAATCCATC 864
QY 598 TTGATGCGCTGAAGAGACACAGATGTGATGTTGGTGGGAACAAGTGTGTGTGT 657
Db 865 CTGACAGCCTCAAGCGCTCCAGCGAGCTGATGTTGGTGGCAAGCAGTGTGTGTGT 924
QY 658 GGCATAGGTGAGGTAGGCAAGGCGTCTGTCTCTCAAGCTCTTGAGCAATGTCTC 717
Db 925 GGCTACGGCGATGTGGGCAAGGATGTCTCAAGCTCTAAAGGGCAAGGCTGCATTTGTG 984
QY 718 TACATTAACCGAAATGACACCCATCTGTCTCTGACAGGCTGCATGATGGTTCAGGCTG 777
Db 985 TACATCAAGAGATGATCCATTTGCGCCCTGACAGGCGAGATGATGCTTCCGTGTG 1044
QY 778 GTAAAGCTAAATGAAGTCAATCCGGCAGTGTGATGTAATTAATCTGCACAGGAATAG 837
Db 1045 GTCAAGCTGAACGAGGTGATCCGCAAGTAGACATTTGTGTGAGCGGCAAGGCAACAG 1104
QY 838 AATGATGACACGGGACGACCTTGTGATGCAAAAAACAGTGTATGATGCAATATG 897
Db 1105 AAGTGTGTGTGCGGAGCAGACATGACAAAGATGAAGAGTGTGCTGATGCAACATG 1164
QY 898 GGCACATCCAAACAGAAATGATGACAGCCTCCGACATCCGAGAGTGTGAGTGGAG 957
Db 1165 GGCACATCGAACAAGGAATGATGATGAGGCTGCGCACACCGGACTTGACCTGGAG 1224
QY 958 CGAGTACGTTCTAGGTGACCATGATCTGAGCCAGATGGCAACAGATGTCTCTCTG 1017
Db 1225 AAGTGGCTCCAGGTGATGATCAATTAATCTGCGGAGGCGCAATCATCATCTTCTG 1284
QY 1018 GCAGAGGCTGCTACTCAATTTGAGCTGCTCCACAGTTCACACTTGTCTGTCTATC 1077
Db 1285 GCCGAGGCGAGGCTGTCAATCTGAGCTGCTCCAGCATCCCTGCTGTGCGCTGTATC 1344
QY 1078 ACAGCCACAACAGAGGCTTTGGCACTGATGAACTCTAATATGACCCGAGGGCGATAC 1137
Db 1345 ACCTGGCCACACAGGCGCTGGCCCTGATGAGCTTTCAATGCCCCCGGAGCGCTAC 1404
QY 1138 AAGCAGATGTGTACTTGTCTTAAGAAATGATGAATAGCTTGCAGCTTGATCTG 1197
Db 1405 AAGTCGATGTCTACTTGTCTGCCAAGAGATGACGAGTATGTGCCAGCTGCACCTG 1464
QY 1198 CCATCAATTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGAGCTC 1257
Db 1465 CCCACCTTCGATGCCCATTTGACGAGCTGAGCGAGCAGGAGGCAAGTACATGGGACTG 1524
QY 1258 AACAAAAATGGGCCATTCAAACTTAATTAATTAACAGATAC 1296
Db 1525 AACCAAGGCGGCTCTTCAAGCCCACTACTACCGCTAC 1563
RESULT 7
ABL29787
ID ABL29787 standard; DNA; 1776 BP.
XX
AC ABL29787;

XX 26-MAR-2002 (first entry)
DT
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 40834.
XX
KM Drosophila; developmental biology; cell signaling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN W0200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
PI
XX
DR WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 40834; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (AB57737-AB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1776 BP; 437 A; 432 C; 479 G; 428 T; 0 other;

Query Match 48.5%; Score 628.4; DB 23; Length 1776;
Best Local Similarity 68.4%; Pred. NO. 9.7e-194;
Matches 885; Conservative 0; Mismatches 406; Indels 3; Gaps 1;

QY 1 TCCAGGCGCAGCAATTTCTGTGAAGACATCAAGCAGGAGATTTGGACGCCG 60
DB 379 TCCGCCGAGGCGACTGACTTTTGGCTGAAGACATCTCGAAGAGTGCATTTGGAGGCCG 438
QY 61 GAGATTGAGATTCAGACAGACATGCTCTGATTTCACTCAGGAAACGTCTCAG 120
DB 439 GAGATCGAGATCCGCCGAGTCGAGATGCCGGCATCATGACTCTGAGGAAAGAGCGAAG 498
QY 121 GGGAGAGAGCCCTTGGCTGTGCTAAATAGTGGCTGTACACACATCACAGCCGACACA 180
DB 499 GATGAGAGAGCCCTTAAAGGCTGCAATATCGTCGATGACACCCACGCTCAATGCTCAGTCG 558
QY 181 GCGGTTGATGAGACACTCTGTGCTTGGGGCTCAGTGCCTGCTGCTTGTAAAC 240
DB 559 GCAGTGTGATCGAGACCTCTGTCACACTGGGCGCCACAGTTCGCTGGGCTGCTGCAAC 618
QY 241 ATCTACTCACTCAGATGAAGTAGCTGACAGACTGGCTGAGGCTGAGTGCAGTGTTC 300
DB 619 ATTATTCACACAAAGCAGATGTCGCCGCTCTGCGACAGAGCGGGAATCCGATCTTC 678
QY 301 GCTTGAAGGCGCAGCTCAGAGATGACTTCTGCTGTGCTATTTGACCCGCTGTGAACATG 360
DB 679 GCCTGGCGCGGAGAGACGGAAGAGAGTCTGTGCTGTGCTTGGACAGGGCCATCTACTCC 738

QY 361 GATGGGTGGCAGGCCAATCATGATCTCTGATGATGGGGAGACTTAACCCACTGGTTAT 420
DB 739 GACGGCTGGCAGCCGAACCTAATCTCTGACGACGGCGGATGCCACGACCTTAATCTC 798
QY 421 AAGAAGTATCCAAAGCTGTTAAGAAGATCCGAGCATTTGGAAGAGAGCGTACTGT 480
DB 799 AAGAAGTATCCCGACTACTTCAAGGCCATTCGGGGCATCTGGAAGAAAGTGTGACCCGG 858
QY 481 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAGGCTGTGTTCGGCCATGACGTC 540
DB 859 GTGCACCGGCTGTATCTGCTGTCAAAGGGCGGAAACTTACTGTTCCGGCCATCAACGTT 918
QY 541 AATGATCTGTTACCAACAGAAAGTTGATTAATGTTGTTGTTGTTGTTGTTGTTGTTG 600
DB 919 AACGACTGATGACCAAGAAAGTTGATTAATGTTGTTGTTGTTGTTGTTGTTGTTGTTG 978
QY 601 GATGGCTGAAGAGACCAAGATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 660
DB 979 GACAGTCTGAAGACCAAGATGATTAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1038
QY 661 TATGTTGAGTAGGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 1039 TACGGTGTATGGGAAGGGCTGTGCCAGTCCCTGAAGGGCCAAAGATGATGATGTTAT 1098
QY 721 ATTACCGAATCGACCCCATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 1099 GTTACGGAAGTGAATCCATATGTGCTCTTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1158
QY 781 AAGCTAATGAATCATCCGCGCAAGTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
DB 1159 CGGCTCAACGAGGCTATCATGAGGACGGGTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1218
QY 841 GTAGTGACAGGGAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
DB 1219 GTTATTTACGAGGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1278
QY 901 CACTTCAACACAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 1279 CATTCCTGCTCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1338
QY 961 GTAGCTTCTCAGTGTGACATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
DB 1339 GTCCGTCTCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1398
QY 1021 GAGGCTGCTACTCTCAATTTGAGCTGCTCCACAGTTCCTTGTCTGCTGCTGCTGCTGCT 1080
DB 1399 GAGGGAAGACTGTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1458
QY 1081 GCCACAACACAGGCTTGGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1459 TCATCCACCCAGGCTTGGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1515
QY 1141 CAGATGTGATCTGCTTCTTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1516 TCGGATGTCTACCTGCTGCAAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1575
QY 1201 TCATTTGATGCCCATCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAAC 1260
DB 1576 ACCTTCGATGCTCATCTCAAGGAGCTCAGCGATGACGACGCTCAAGTTATGGGCTTAAAC 1635
QY 1261 AAAAATGGCCATCAAACTAATTTATACAGAT 1294
DB 1636 AAGCGCGGCTTTTAAAGCCAAATTACTACAGGT 1669

RESULT 8
ABL03692/c
ID ABL03692 standard; cDNA; 4406 BP.
XX
AC ABL03692;
XX
DT 26-MAR-2002 (first entry)
XX

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 5558.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PMD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR P-PSDB; ABB59589.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 5558; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL1840-ABL16175), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SO Sequence 4406 BP; 1112 A; 1039 C; 1064 G; 1191 T; 0 other;

Query Match 37.5%; Score 486.2; DB 23; Length 4406;
Best Local Similarity 66.9%; Pred. NO. 4.1e-147;
Matches 742; Conservative 0; Mismatches 303; Indels 64; Gaps 1;

QY 252 TCAGATGAGTACCTGACGACTGCTGAGGCTGAGTGCAGTGTGCTGGAAGG 311
DB 2112 TTAGACGAGTGTGCTGCTGCTTGGCCGAATCCGATCTTGTGCTGCGCGG 2053

QY 312 CGAGTCAGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 371
DB 2052 CGAGACGAGAGGACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1993

QY 372 GGCCACATGATCCTGATGATGGGAGAGACTTAACCACTGGTTTATAAGATATCC 431
DB 1992 GCCAATCATGATCCTGAGCATGCGGCGGATGCCACGCACTGATGCTAAAAAGTACC 1933

QY 432 AAAGCTGTTAAGATCCGAGGATGCTGGAAGAGAGAGAGAGAGAGAGAGAGAG 491
DB 1932 GAAGATGTTCAAGTTGCTCAAGGAATGCTGAGAGAGAGAGAGAGAGAGAGAG 1873

QY 492 GTATCAGCTCTCCAAAGCTGGGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 551
DB 1872 CTATCAGCTCTCTAAGCGCGCAAGCTGACAGTGCAGGATGTAATGTAACGATTCGGT 1813

QY 552 TACCAACAGAGTTGATTAATCTGTACTGCTGCGGAGAAATCCATTTTGATGCGCTGAA 611
DB 1812 GACAAAGACCAATTTGACAACCTCTACAGCTGCAAGAAATCCATCTGACAGCCCTCAA 1753

QY 612 GAGGACACAGATGATGTTGTTGGTGAACCAAGTGTGTGTGTGTGTGTGTGTGTGT 671
DB 612 GAGGACACAGATGATGTTGTTGGTGAACCAAGTGTGTGTGTGTGTGTGTGTGTGT

DB 1752 GCGCTCCACGAGCGTGAATGTTCCGTTGGCAACGAGGCTGCTGTGTGCTACGGCATGT 1693
QY 672 AGGCAAGGGCTGCTGTGCTCTCA----- 697
DB 1692 GGGCAAGGATGTGCTCAGGCTCTAAAGGGGAGGTAATGGCACCACCAACCATAGTA 1633
QY 698 -----AAGCTCTGGAGCAATTGCTACATTAACG 727
DB 1632 TGATATACCTTTGTACCTTTTCAATCTCCAATCGTAGGGCTGCATTTGTGATACACAG 1573
QY 728 AATCGACCCCATCTGTGCTCTGACGGCTGCATGGATGGTTCAAGGTGTAAGCTAA 787
DB 1572 AGATCGATCCCATTTGCGCCCTGCAGGCGCATGATGGCTTCCGTGTGCTAAGCTGA 1513
QY 788 ATGAAGTCATCCGGCAGTGCATGTCGTAATTAATCTGCACAGAAATAGATGTAGTGA 847
DB 1512 ACGAGGTGATCCGCAACGTAGACATTGTGTGACGGCGGACCGCAACAGAACGTGTGG 1453
QY 848 CACGGGAGCATTGATCGCATGAAAAACAGTTGATGATGCAATATGGCCACTCCA 907
DB 1452 TCGCGAGCAGCATGACAGATGAAGAGTGGCTGCAATGATGCAACATGGCCACTCGA 1393
QY 908 ACACAGAAATCGATGTGACAGCCTCCGCACTCCGAGCTGACGTGGAGCGATACGTT 967
DB 1392 ACACGGAATCGATGTGATGAGCTTCGACACACCGGACTTGACCTGGAGAGAGTGGCT 1333
QY 968 CTCAGGTGACCATGTCTCTGCGCCAGATGGCAACAGTGTCTCTCTGCGAGAGGTC 1027
DB 1332 CCCAGGTGATCATATATCTGCGCGGAGGCAAGTACATCATCTTCTGCGCGAGGCA 1273
QY 1028 GTTACTCATTTGAGCTGTGCTCCACAGTTCACACCTTTGTTCTGCTACACAGCAAA 1087
DB 1272 GGCTGTCAATCTGAGCTGTCTCCAGCATCCCTCGTGTGCGCTGTCTATCACCTGGCCA 1213
QY 1088 CACAGGCTTGGCAGCTGATGAACCTCTAATATGACCCGAGGGCGGATACAGCAGATG 1147
DB 1212 CCCAGGCTGTGCGCTGATGAGCTTTTCAATGCCCCGCGGACGCTACAGTGGATG 1153
QY 1148 TGTACTGTCTTCTAAGAAATGATGATGATGATGATGATGATGATGATGATGATG 1207
DB 1152 TCTACTGTGCTGCGCAAGATGAGAGATGATGATGATGATGATGATGATGATGATG 1093
QY 1208 ATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGACTCAACAAAATG 1267
DB 1092 ATGCCATTTGACGAGCTGAGCGACGAGCAGGCAAGTACATGGAGTGAACAGGCG 1033
QY 1268 GGCCATTCAAACTAATTAATACAGATAC 1296
DB 1032 GTCCCTTCAAGCCCAACTACTACCGCTAC 1004

RESULT 9
ABL79045
ID ABL79045 standard; cDNA; 553 BP.
XX
AC ABL79045;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human ovarian cancer related cDNA clone SEQ ID NO:2023.
XX
KW Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200192581-A2.
XX
PD 06-DEC-2001.
XX
PF 29-MAY-2001; 2001WO-US17756.
XX
PR 26-MAY-2000; 2000US-207484P.
XX

PA (CORI-) CORIXA CORP.
XX
PI Algate PA, Harlocker SL, Jones R;
XX
DR WPI; 2002-122075/16.
XX
PT Composition for therapy and diagnosis of ovarian cancer comprising
PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
PT polypeptide, antibody specific to polypeptide or T cell expressing
PT polypeptide

PS Claim 1; SEQ ID 2023; 489pp; English.

XX
CC The present invention describes a composition (I) comprising: carriers
CC and immunostimulants; and a polypeptide (II) of a ovarian tumor
CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence
CC (S1) from the 10912 nucleotide sequences as given in ABL77023 to
CC ABL87934, (III) encoding (II) having a sequence (S2), a T cell
CC population of (II), or antigen presenting cells that express (II).
CC (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to
CC (S1) can be used for detecting ovarian cancer in a patient's biological
CC sample preferably serum or ovarian tissue. The method comprises
CC contacting a biological sample from a patient with (IV), detecting the
CC amount of polynucleotide hybridising to (IV) and comparing the amount to
CC a predetermined cutoff value and thereby detecting ovarian cancer in the
CC patient, where the amount of polynucleotide hybridising to (IV) is
CC detected preferably by polymerase chain reaction (PCR). (I) comprising
CC (III) and/or (II) is useful for stimulating and/or expanding T cells
CC specific for an ovarian tumor protein comprising and contacting T cells
CC with (III) or (II). (III) is useful in design and preparation of
CC ribozyme molecules for inhibiting expression of the tumor polypeptides
CC and proteins in tumor cells; and to isolate a full length gene from a
CC suitable library e.g., a tumor CDNA library using well known
CC techniques.

XX
SQ Sequence 553 BP; 159 A; 136 C; 128 G; 127 T; 3 other;

Query Match 35.6%; Score 461.4; DB 24; Length 553;
Best Local Similarity 97.3%; Pred. No. 1.4e-139;
Matches 510; Conservative 0; Mismatches 9; Indels 5; Gaps 4;

OY 774 GGTGTTAAAGCTAAATGATGATCCGGCAAGTCGTAATTAATGACACAGGAA 833
DB 1 GGTGTTAAAGCTAAATGATGATCCGGCAAGTCGTAATTAATGACACAGGAA 60
OY 834 TAAGATGTAGTACACGAGCACTTGATCGCATGAAGAAACAGTTGATCGTACAA 893
DB 61 TAAGATGTAGTACACGAGCACTTGATCGCATGAAGAAACAGTTGATCGTACAA 120
OY 894 TATGGGCACTCCACACAGAAATCGATGTGACACAGCCTCCGCACTCCGAGCTGACGTG 953
DB 121 TATGGGCACTCCACACAGAAATCGATGTGACACAGCCTCCGCACTCCGAGCTGACGTG 180
OY 954 GGAGCGAGTACGTTCTCAGGTGGACCATGTCTATGCGCAGATGGCAACGAGTGTCT 1013
DB 181 GGAGCGAGTACGTTCTCAGGTGGACCATGTCTATGCGCAGAT-GNAAACGAGTGTCTCT 239
OY 1014 CCTGGCAGAGGGTCTCTACTCAATTGAGCTGCTCCACAGTCCACCTTGTCTGTC 1073
DB 240 CCTGGCAGAGGGTCTCTACTCAATTGAGCTGCTCCACAGTCCACCTTGTCTGTC 299
OY 1074 CATGACAGCCACACACAGGCTTGGCAGCTGATAGACTCTATATGACCCGAGGGCG 1133
DB 300 CATGACAGCCACACACAGGCTTGGCAGCTGATAGACTCTATATGACCCGAGGGCG 359
OY 1134 ATACACAGCAGATGTGTACTGCTTCTAAGAAA-TGGATGAATACGTTGCCAGCTGC 1192
DB 360 ATACACAGCAGATGTGTACTGCTTCTAAGAAAATGATGAATACGTTGCCAGCTGC 419
OY 1193 ATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1252
DB 420 ATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAGC-AAATATCTGG 478

OY 1253 GACTCACAAAATGGCCATTCACACCTAATTATACAGATAC 1296.
DB 479 GACTCACAAA--TGGCCATTCACACCTAATTATACAGATAC 520

RESULT 10

ABV44141/C
ID ABV44141 standard; CDNA; 504 BP.

XX AC ABV44141;

DT 16-SEP-2002 (first entry)

XX DE Human prostate expression marker CDNA 44132.

XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.

XX OS Homo.sapiens.

XX PN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US05171.

XX PR 17-FEB-2000; 2000US-183319P.

XX PR 16-MAR-2000; 2000US-189862P.

XX PR 25-MAY-2000; 2000US-207454P.

XX PR 09-JUN-2000; 2000US-211314P.

XX PR 18-JUL-2000; 2000US-219007P.

XX PR 13-DEC-2000; 2000US-255281P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE;

XX DR WPI; 2001-662795/76.

XX PT Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer; useful
XX for detecting presence of prostate cancer, stage of prostate cancer -
XX Claim 1; Page 8770; 11750pp; English.

XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for:
XX (a) assessing whether a patient is afflicted with prostate cancer;
XX (b) monitoring the progression of prostate cancer in a patient;
XX (c) assessing the efficacy of a test compound to inhibit prostate
XX cancer in a patient;
XX (d) assessing the efficacy of a therapy for inhibiting prostate cancer
XX in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound;
XX (g) determining whether prostate cancer has metastasized in a patient;
XX (h) assessing the aggressiveness or indolence of prostate cancer in a
XX patient;
XX (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX SQ Sequence 504 BP; 117 A; 150 C; 116 G; 121 T; 0 other;

Query Match 32.6%; Score 422; DB 23; Length 504;
Best Local Similarity 100.0%; Pred. No. 9.2e-127;
Matches 422; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 158 GTACACATCACAGCCACAGACGCGGTGTGATGAGACACTGTGCCCCGCGGCTC 217
DB 502 GTACACATCACAGCCACAGACGCGGTGTGATGAGACACTGTGCCCCGCGGCTC 443
OY 218 AGTGGCGGTGTGCTGTGTAACATCTACTCACTCAGAAATGAAGTAGCTGACAGACTGG 277

DB 442 AGTCCGCTGCTGCTTGTACATCTACTCACTCAGAAATGAGTAGCTGCAGCACTGG 383
QY 278 CTGAGGCTGAGTTCAGTGTTCCTTGGAGGCGAGTCAGAGATGACTTCTGTGCT 337
DB 382 CTGAGGCTGAGTTCAGTGTTCCTTGGAGGCGAGTCAGAGATGACTTCTGTGCT 323
QY 338 GTATGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 397
DB 322 GTATGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 263
QY 398 GAGACTTAACCCACTGCTTATTAAGAGTATCCAAAGCTGTTAAGAGATCCGAGCA 457
DB 262 GAGACTTAACCCACTGCTTATTAAGAGTATCCAAAGCTGTTAAGAGATCCGAGCA 203
QY 458 TTGTGAGAGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 517
DB 202 TTGTGAGAGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 143
QY 518 TCTGTGTTCCGCGCAGTCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 577
DB 142 TCTGTGTTCCGCGCAGTCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 83
QY 578 AC 579
DB 82 AC 81

RESULT 11

ABL29786 ID ABL29786 standard; DNA; 5537 BP.

AC ABL29786;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 40831.

KW Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ds.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions

PS Claim 1: SEQ ID NO 40831; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB16176-AB130511), expressed DNA sequences (AB101840-AB16175) and the encoded proteins (AB157737-AB172072).
CC The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 5537 BP; 1632 A; 1054 C; 1158 G; 1693 T; 0 other;

QY Query Match 31.4%; Score 406.6; DB 23; Length 5537;

Best Local Similarity 60.8%; Pred. No. 4.3e-121;

Matches 801; Conservative 0; Mismatches 379; Indels 137; Gaps 3;

QY 1 TCCAGGCGAGCAGCAATTTCTGTGTAAGAACATCAAGCAGCAGCAATTTGAGCGCG 60
DB 2937 TCCGCGGAGGCGACTGACTTTGCGTGAAGAGCATCTCGAAGAGTGCATTTGAGCGCG 2996
QY 61 GAGATTGAGATTGCAGAGCAGACATGCTGCTGATTTCACTCAGGAAACGTGCTCAG 120
DB 2997 GAGATCGAGATCGCCGAGTCGAGATGCCGGCATCATGACTCTGAGGAAAGAGCGAAG 3056
QY 121 GGGGAGAGCCCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
DB 3057 GATGAGAGAGCCCTTAAGGGTGCCTAATATGCTGCTGCTGCTGCTGCTGCTGCTGCT 3116
QY 181 GCGGTGTTGATTGAGACACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
DB 3117 GCACTGCTGATCGAGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3176
QY 241 ATCTACTCAACTCAGAAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 300
DB 3177 ATTTATTCACACAAACGACAGTTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3236
QY 301 GCTTGAAGGCGAGTCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 360
DB 3237 GCGTGGCGGAGAGAGACGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3296
QY 361 GATGGTGGCAGGCGCAACATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB 3297 GACGGCTGGCAGCGCAACCTAATCTGAGCAGCAGCGCGCGGATGCCACGACCTATGCTC 3356
QY 421 AAGAGTATCCAAACGCTTTAAGAAATCCGAGGATTTGGAAGAGAGAGAGAGAGAGAGAG 480
DB 3357 AAGAGTATCCCGACTACTTCAAGGCTTGGGGCATCTGAGAGAGAGAGAGAGAGAGAG 3416
QY 481 GTTCAGGCTGTATCAGCTCTCCAAAGCTGGGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 540
DB 3417 GTGACCGGCTGTATCATGCTGCTCAAGGCGGAAACTTACTGCTGCTGCTGCTGCTGCTGCT 3476
QY 541 AATGATCTGTATCCAAACAGAGTGTGATACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
DB 3477 AAGACTCAGTGACCAAGAAAGATTTGATCTTCTACAGCTGCTGCTGCTGCTGCTGCTGCT 3536
QY 601 GA----- 602
DB 3537 GACAGGTGAGCATAGATCTCATGATTAAGGAAGTACTATCTGCTGCTGCTGCTGCTGCT 3596
QY 603 -----TGGCTGAAGAGAGCAGCAGATGTGATGTTGGTGGGAAACAAGTGG 649
DB 3597 AATAATCTCTTCCAGTCTGAACGACACCGATATATGTTGGCGGAAAGCAGGTGG 3656
QY 650 TGGTGTGGCTATGCTGAGTGAAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 709
DB 3657 TGATCTGTGGTACGCTGATGTGGAAGGCTGTGCCAGTCCCTGAAGGCGCAAGGAT 3716
QY 710 CAATTGCTACATTAACGAAATCGACCCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 769
DB 3717 GCATGTTTATGTTACGGAAGTGAATCCATATGCTGCTTACAGCTGCCATGATGATGAT 3776
QY 770 TCAGGCTGTAAGCTAATGAAGTCAACGCGCAAGTGCATGCTGCTGCTGCTGCTGCTGCTGCT 829
DB 3777 TCCGGTGTGAGCGCTCAACGAGGTCATCAGAGCGGTGATGCTGCTGCTGCTGCTGCTGCTGCT 3836
QY 830 GAATAAGATGTAGTACACGAGGAGCACTTGATGCGATGAAGAAACAGTTGATGCTGAT 889
DB 3837 GAACAAATATGTATTACAGGAGTACATGATGATGATGATGATGATGATGATGATGATGAT 3896
QY 890 GCAATATGGGCACTCCACACAGAAATCGATGT----- 923

DB 3897 GCAATATGGACATCTCTGCTGGAGATGTGATGTATGTATTCGGGGTGAATTCATTG 3956
QY 924 -----GACCAGCCTCCGCACTCCGGAGC 946
DB 3957 TGGCCTTCATATATCAAAAGTTTAAACCACTTTAGAATGGCTTGCAATACCCCGAGC 4016
QY 947 TGACGTGGAGCAGATAGCTTCTCAGGTGACCATGTCATCTGGCCAGATGGCAACGAG 1006
DB 4017 TAACGTGGAGCGGTGCTCTCTCAAGTGCATCATCAGGTGGCCGAGCGAGATGA 4076
QY 1007 TTGTCCTCCTGGCAGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTTCACCTTTG 1066
DB 4077 TCATTTGCTCGCCGAGGGAAGACTGGTAATTTGCTCTGTTCCACCATTTCTCTTTG 4136
QY 1067 TTCTGTCCATCACAGCACAACAGGCTTTGGCACTGATAGAACTCTATATGCAACCCG 1126
DB 4137 TCGTATCCGTGGCTCATCCACCCAGGCTTTGGCCCTGATTTGAACCTCTTCTCAAGCGCC - 4194
QY 1127 AGGGCGATACAGCAGGATGTGTACTTGTCTTAAGAAATGGATGAATACGTTG 1183
DB 4195 -AGGAAGATATAGTCGATGTCTACCTGTGCCAAGAAATGGTAGGTTCTTAG 4250

RESULT 12
AAZ80766/c
ID AAZ80766 standard; cDNA; 636 BP.

XX AAZ80766;

DT 07-APR-2000 (first entry)

XX Human colon cancer cell line SW480 cDNA clone SEQ ID NO:850.

DE Human; gene expression product; diagnosis; tumour; colon cancer;

KW colorectal adenocarcinoma; cell line SW480; cell proliferation;

KW cytostatic; sarcoma; breast cancer; neoplasia; dysplasia;

KW hyperplasia; ds.

XX Homo sapiens.

XX WO9964576-A2.

XX 16-DEC-1999.

XX 09-JUN-1999; 99WO-IB01062.

XX 10-JUN-1998; 98US-0088801.

PA (FARB) BAYER CORP.

XX Endege WO, Steilmann KE, Astle JH, Burgess CC, Bushnell SE;

PI Carroll E, Catino TJ, Dertl A, Ford DM, Lewis ME, Monahan JE;

PI Schlegel R;

XX WPI; 2000-087220/07.

XX Novel nucleic acids, used to develop products for the diagnosis and

PT treatment of disorders involving unwanted cell proliferation,

PT particularly cancers, especially colon cancer

XX Claim 15; Page 469; 469pp; English.

XX AAZ79917 to AAZ80766 represent double stranded cDNA clones isolated from
CC the human colorectal adenocarcinoma (colon cancer) cell line SW480. The
CC cDNA clones can be used to generate antisense oligonucleotides which
CC can be used for antisense therapy. Methods and products from the present
CC invention can be used for identifying and/or classifying cancerous cells
CC present in a human tumour, particularly in solid tumours, e.g.
CC carcinomas and sarcomas, e.g. breast or colon cancers. The cDNA clones
CC can be used for developing agents for the diagnosis and treatment of
CC disorders involving unwanted cell proliferation, such as neoplasia,
CC dysplasia or hyperplasia.

XX SQ Sequence 636 BP; 156 A; 166 C; 137 G; 147 T; 30 other;

Query Match 31.4%; Score 406.4; DB 21; Length 636;
Best Local Similarity 99.5%; Pred. NO. 1.3e-121;
Matches: 418; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 159 TACACACATCACAGCCAGACAGC-GGTGTGATTTGAGACACTCTGTGCCCTGGGGCTC 217

DB 420 TACACACATCACAGCCAGACAGCGGGTGTGATTTGAGACACTCTGTGCTGGGGCTC 361

QY 218 AGTCCCGTGTCTGTGTGATACATCTACTCACTCAGATGAGTAGCTGCAGCACTGG 277

DB 360 AGTCCCGTGTCTGTGTGATACATCTACTCACTCAGATGAGTAGCTGCAGCACTGG 301

QY 278 CTGAGGCTGAGTTGCACTGTTCCCTTGGAGGGCCGAGTCAGAGATGACTTCTGGTGT 337

DB 300 CTGAGGCTGAGTTGCACTGTTCCCTTGGAGGGCCGAGTCAGAGATGACTTCTGGTGT 241

QY 338 GTATTGACCGCTGTGTGACATGATGGTGGCAGGCCAATGATCTCTGATGATGGG 397

DB 240 GTATTGACCGCTGTGTGACATGATGGTGGCAGGCCAATGATCTCTGATGATGGG 181

QY 398 GAGACTTAACCCACTGGGTTTATAGAGATATCCAAAGTGTTTAGAGATCCGAGCA 457

DB 180 GAGACTTAACCCACTGGGTTTATAGAGATATCCAAAGTGTTTAGAGATCCGAGCA 121

QY 458 TTGTGAGAGAGAGCGTGTGTTTACAGAGCGTGTATCAGCTCTCCAAAGCTGGGAGC 517

DB 120 TTGTGAGAGAGAGCGTGTGTTTACAGAGCGTGTATCAGCTCTCCAAAGCTGGGAGC 61

QY 518 TCTGTGTTCCGGCCATGACGTCAATGATCTGTACCAACAGAGTTGATACTGT 577

DB 60 TCTGTGTTCCGGCCATGACGTCAATGATCTGTGTACCAACAGAGTTGATACTGT 1

RESULT 13
AAC77823
ID AAC77823 standard; cDNA; 2200 BP.

XX AAC77823;

DT 08-FEB-2001 (first entry)

XX Human cancer associated gene sequence SEQ ID NO:217.

XX Human; cancer associated gene; cancer antigen; detection; cancer;

KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;

KW antidiabetic; antidiabetic; antirheumatic; antirheumatic; antiviral;

KW antidiabetic; antidiabetic; antirheumatic; antirheumatic; antiviral;

KW dermatological; neuroprotective; thrombolytic; coagulant; neotropic;

KW vasotropic; antipsoratic; antiangiogenic; gene therapy; inflammation;

KW immune disorder; haematopoietic cell disorder; autoimmune disorder;

KW allergic reaction; graft versus host disease; organ rejection;

KW haemostatic; thrombolytic; cardiovascular disorder; infection;

KW neurological disease; drug screening; ss.

XX Homo sapiens.

XX WO200055350-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05882.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM;

XX WPI; 2000-587533/55.

XX P-PSDB; AAB43614.

xx Novel isolated nucleic acids comprising sequences encoding peptides.
 PT useful for treating or diagnosing e.g. cancer -
 xx Claim 1; page 787-788; 2352pp; English.

xx AAC77607 to AAC78448 encode the human cancer associated proteins given
 CC in AAB4398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnerary; immunomodulator;
 CC antidiabetic; antiasthmatic; antirheumatic; antiallergic;
 CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
 CC dermatologic; neuroprotective; cardiant; thrombolytic; coagulant;
 CC nootropic; vasotropic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention.

xx Sequence 2200 BP; 546 A; 578 C; 607 G; 466 T; 3 other;

Query Match 31.0%; Score 402.4; DB 21; Length 2200;
 Best Local Similarity 58.4%; Pred. No. 5.7e-120;
 Matches 741; Conservative 0; Mismatches 521; Indels 6; Gaps 2;

OY 32 ACATCAAGCAGGCGAATTTGGACCGCGGAGATTGAGATTGACAGCAAGACATGCTCG 91
 DB 108 ACATCGGCTGGCTGGCTGGGAGCGGACTCGGCTCCAGCCACTGAAGGGCGCCGATCG 167
 OY 92 CTCTGATTTCACTCAGGAACGCTGCTCAGGGGAGAGAGCCCTTGCTGCTAAATAG 151
 DB 168 GCCTGATGCGTATGCGGAGCGGACTCGGCTCCAGCCACTGAAGGGCGCCGATCG 227
 OY 152 TGGGCTGTACACACATCACAGCCAGACAGCGGTGTGATGAGACACTCTGCCCCG 211
 DB 228 CTGGCTGCTGCACATGACCGGTGAGAGAGCGGCGCTCTCATTTGAGACCTCTGCA 287
 OY 212 GGGCTCAGTCCGCTGCTGCTGCTGTAACATCTACTCACTCAAGTGAAGTACGCTG 271
 DB 288 GTGCTGAGGTGAGTGTCCAGCTGCAACATCTTCTCCACCCAGGACCATGCGGCGCTG 347
 OY 272 CACTGGCTGAGGCTGAGTTGCGTGTTCGCTTGAAGGGCGAGTCAAGATGACTTCT 331
 DB 348 CCAATTGCCAAGGCTGCGCATTCGCTGTATGCTGGAAGGGCGAAGCGAGAGATACC 407
 OY 332 GGTGGTGTATTGACCGCTGTGTGAACATGATGGTGGCAGGCCAATATGATCTGGATG 391
 DB 408 TGTGGTGCATGTAGCAGACCCCTTACTTCAAGAGAGCGGCGCCCTCAACATGATCTG 467
 OY 392 ATGGGGGAGACTTAACCCACTGGGTTTATAAGAGTATCCAAAGCTGTTAAGAGATCC 451
 DB 468 ACCGGGGGCGACCTCACCAACCTTATCCACCAAGTACCCGCACTTCTGCCAGGATCC 527
 OY 452 GAGGCAATTGTGAAGAGAGCGGTGCTGCTTCAACAGGCTGTATCAGCTCTCCAAAGCTG 511
 DB 528 GAGGCATCTCTGAGGAGACCACTGGGGTCCACAACTCTTACAGATGATGGCCAATG 587
 OY 512 GGAAGCTCTGTGTTCCGGCCATGAACGTCAATGATTTGTTACCAAAACAGAGTTTGATA 571
 DB 588 GGATCCTCAAGGTGCTGCTGCAATGATGATGACTCCGTACCAAGAGCAAGTTTGACA 647
 OY 572 ACTGTACTGCTGCGGAGAAATCAATTTTGATGGCTGAAGAGAGCAACAGATGTGATGT 631
 DB 648 ACCCTATGCTGCGGAGAGTCCCTCATAGATGATCAAGCGGGCCACAGATGTGATGA 707

OY 632 TTGGTGGAAACAAGTGTGTGTGTGCTATGTAGGTAGGCAAGGCTGTGCTG 691
 DB 708 TTGCGGCAAGGTAGCGGTGTGTAGCAGAGGTATGTGTGTGGGCAAGGCTGTGCCAGG 767
 OY 692 CTCTCAAGCTCTTGAGCAATTTGCTTACATTAACGAAATCGACCCCACTGTGCTGTC 751
 DB 768 CCCTGCGGGTTTCGAGACCCGCGCTCATCATCAACGAGATTGACCCCAACGCACTGC 827
 OY 752 AGGCTGCATGATGGGTTTCAAGGTGTGTAAGCTAAATGAAGTCAATCGGCAAGTGATG 811
 DB 828 AGGCTGCATGATGGGTTTCAAGGTGTGTAAGCTAAATGAAGTCAATCGGCAAGTGATG 887
 OY 812 TCGTATACTTGCACAGGAATAAGATGTGACACAGGAGCACTTGGATCGCATGA 871
 DB 888 TCTTGTACACCAACAGAGCTGTATTTGACATCATCTTGGCCGCACTTTGAGCAAGATGA 947
 OY 872 AAACAGTTGTATGCTATGCAATATGGGCCCACTCCAAACAGAAATCGATGTGACAGCC 931
 DB 948 AGGATGATGCCATTTGTGTGTATACATTTGACACTTTGACGTGAGATGATGCAAGTGCC 1007
 OY 932 TCCGCACTCCGAGCTGACGTGGGAGCGAGTACGTTCTCAGGTGGACCATGTCATCTGCC 991
 DB 1008 TCAACGAGAACCGCGTGAGAGAGGTGAACATCAAGCCGAGGTGAGCCGATTCGGTTGA 1067
 OY 992 CAGATGCAAAAGATTGTCTCTCTGCGAGAGGCTGCTACTCAATTTGAGCTGCTCCA 1051
 DB 1068 AGAATGGGCGCGCATCATCTCTGCTGCGCGAGGCTGCTGTAACCTGGGTTGTGCCA 1127
 OY 1052 CAG--TTCCACCTTTGTCTGTCTCATCAACAGCCACACAGAGGCTTGGCACTGATAG 1108
 DB 1128 TGGGCCACCCAGCTTCTGTATGATGATGATGATGATGATGATGATGATGATGATG 1187
 OY 1109 AACTCTATAATGACACCCGAGGGCGGATACACAGGATGTGTACTTGTCTTCTTAAGAAA 1168
 DB 1188 AGCTGTGACCCATCCAGACAAG--TACCCTGTGGGGTTCATTTCTGCCCCAAGAAGC 1244
 OY 1169 TGGATGAATAGCTTGCAGCTTGCATCTGCCATCATTTGATGCCCATTAAGAGCTGA 1228
 DB 1245 TGGATGAGGAGTGGCTGAAGCCCACTGGGCAAGCTGAATGTGAAGTGAACCAAGCTAA 1304
 OY 1229 CAGATGACCAAGCAAAATATCTGGGACTCAACAAATGGCCATTAACCTAATTAT 1288
 DB 1305 CTGAGAAGCAAGCCAGTACCTGGGCAATGCTGTGATGGCCCTTCAAGCCGGATCACT 1364
 OY 1289 ACAGATAC 1296
 DB 1365 ACCGCTAC 1372

RESULT 14
 ABV23126
 ID ABV23126 standard; cDNA; 2658 BP.
 XX
 AC ABV23126;
 XX
 DT 16-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 23117.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 OS Homo sapiens.
 XX
 PN W0200160860-A2.
 XX
 XX 23-AUG-2001.
 PD
 PF 20-FEB-2001; 2001WO-US05171.
 XX
 XX 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.

PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PI Schlegel R, Endege WO, Monahan JE;
XX WPI: 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer
XX
PS Claim 1; page 4141-4142; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 2658 BP; 624 A; 727 C; 730 G; 552 T; 25 other;
XX
Query Match 31.0%; Score 402.4; DB 23; Length 2658;
Best Local Similarity 58.4%; Pred. No. 6.4e-120;
Matches 741; Conservative 0; Mismatches 521; Indels 6; Gaps 2;
QY 32 ACATCAAGCAGGAGGAAATTTGGACGCCGGAGATTGAGATTGAGAGCAAGACATGTCG 91
DB 355 ACATCGGCGCTGGCTGGCGGAGCGCAAGGCCCTGGACATTCGTGAGAACGAGATGCCGG 414
QY 92 CTCTGATTTCACTCAGGAAAGCTGCTCAGGGGAGAGAGCCCTTGCTGCTAAATAG 151
DB 415 GCTGATGCGTATGCGGAGCGGTACTCGGCTCCAGGCACTGAAGGCGCGCATCG 474
QY 152 TGGGCTGTACACATACACAGCCAGACAGCGGTGTTGATTGAGACACTCTGTCCTCG 211
DB 475 CTGGCTGCTGCACATGACCGGTGAGAGCGCGCTCTCATTTGAGACCTCGTACCCTCG 534
QY 212 GGGCTCAGTGCCTGCTGCTGTTGTAACATCTACTCACTCACTCAGATGAGTAGTCAG 271
DB 535 GTGCTGAGGTGAGTGGTCCAGCTGCAACATCTCTCCACCAGGACCATGCGGCGCTG 594
QY 272 CACTGGCTGAGGCTGAGTTGCACTGTTGCTTGAAGGGGAGTCAAGAGATGACTTCT 331
DB 595 CCATTGCCAAGGCTGGCATTCGGGTATGCTGGAAGGGGGAACGAGAGAGTAGTACC 654
QY 332 GGTGCTGATTGACCGCTGCTGTAACATGATGGTGGCAGGCCAAGCATGATCTGGATG 391
DB 655 TGTGTGCTGATTGAGCAGACCCCTGTAATTCAAGGAGCGGCCCTCAACATGATCTGAGC 714
QY 392 ATGGGGAGAGCTTAACCACTGAGTTTATAGAAGTATCAACAGTGTTTAAGAAGATCC 451
DB 715 ACGGGGGGAGCTCAACCACTCACTCAACACAGTACCGGAGCTTCTGCCAGGCACTCC 774
QY 452 GAGGCATTTGGAAGAGAGCGTGAAGTGTTCACAGGCTGTATCAGCTCTCCAAAGCTG 511
DB 775 GAGGCATCTCTGAGAGAGACCAAGTGGGGTCCACCACTCTCAAGATGATGGCCAATG 834
QY 512 GGAAGCTCTGTGTCGGGCGCATGAGCTCATGATTTCTTTACCAACAGAGATTGATA 571
DB 571 GGAAGCTCTGTGTCGGGCGCATGAGCTCATGATTTCTTTACCAACAGAGATTGATA 571

DB 835 GGATCCTCAAGGTGCGTCCATCATGTCAATGACTCCGTCCACAGAGCAAGTTGACA 894
QY 572 ACTGTACTGCTGCCGAGAAATCCATTTTGATGGCCTGAGAGACACAGATGTGATGT 631
DB 895 ACCCTATGCTGCGCGGAGTCCCTCATAGATGGCATCAAGCGGCGCACAGATGTGATGA 954
QY 632 TTGGTGGGAACAAAGTGGTGTGTGCTATGCTGAGGTAGGCAAGGGCTGTGTGTG 691
DB 955 TTGCGGGCAAGGTAGCGGTGTGTAGCAGGCTATGTTGATGTGGGCAAGGGCTGTGCCAGG 1014
QY 692 CTCTCAAGCTCTTGAGACCAATTTCTACATTAACCAATGACCCCATCTGTGCTGTGC 751
DB 1015 CCTGCGGGGTTTGGAGCCCGCTCATCATCAACGAGATTGACCCCATCAACGCACTGC 1074
QY 752 AGGCTCATGATGATGGTTCAAGGTGTTAAAGCTAAATGAGTCAATCCGGCAAGTCATG 811
DB 1075 AGGCTCATGATGATGGTTCAAGGTGTTAAAGCTAAATGAGTCAATCCGGCAAGTCATG 811
QY 812 TCGTAATACTTGACACAGAAATGAATGATGACACAGGAGCACTTGATGCAATGCA 871
DB 1135 TCTTTGTCACACACAGCGGTGATGATCATCATCTTGCGCGCACTTTGAGCAGATGA 1194
QY 872 AAACAGTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 931
DB 1195 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1254
QY 932 TCCGCACTCCGAGCTGACGTGGGAGCGAGTACGTTCTCAGTGGACCATGTATCTGGC 991
DB 1255 TCAACGAGAACCGCGGTGAGAGAGTGAACATCAAGCCGAGTGGACCGGTATCGGTGA 1314
QY 992 CAGATGGCAAAAGAGTGTCTCTCTGCGCAGAGGGTCTCTACTCAATTTGAGCTGCTCCA 1051
DB 1315 AGAATGGGGCGCGCATCATCTGCTGCGCGAGGGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1374
QY 1052 CAG---TTCCACCTTTGTTCTGTCATCAACAGCCACACAGAGCTTTGGCACTGATAG 1108
DB 1375 TGGGCAACCCAGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1434
QY 1109 AACTCTATAATGACCCGAGGGGCGATACAAAGCAGATGTTGATGTTGCTTCTTAAGAAA 1168
DB 1435 AGCTGTGACCCATCCAGACAAG---TACCCGTTGGGTTCAATTCCTGCCAAGAAGC 1491
QY 1169 TGAATGATAGTTGACAGCTTGCATCTGCACTCATTTGATGGCCACCTTACAGAGCTGA 1228
DB 1492 TGAATGATAGTTGACAGCTTGCATCTGCACTCATTTGATGGCCACCTTACAGAGCTGA 1551
QY 1229 CAGATGACCAAGCAAAATATCTGGGACTCAACAAAATGGGCCATTCAAACCTAATATT 1288
DB 1552 CTGAGAGCAAGCCAGTACCTGGGATGTTCTGTGATGGGCCCTTCAAGCCGATCACT 1611
QY 1289 ACAGATAC 1296
DB 1612 ACCGCTAC 1619
RESULT 15
ABV28967
ID ABV28967 standard; cDNA; 2658 BP.
XX
XX AC ABV28967;
XX AC
XX DT 16-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 28958.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX OS
XX PN WO200160860-A2.
XX
XX PD 23-AUG-2001.

```

XX 20-FEB-2001; 2001WO-US05171.
PF
XX
PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer
XX
XX
PS Claim 1; Page 6122; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 2658 BP; 624 A; 727 C; 730 G; 552 T; 25 other;
XX
XX
Query Match 31.0%; Score 402.4; DB 23; Length 2658;
Best Local Similarity 58.4%; Pred. No. 6.4e-120;
Matches 741; Conservative 0; Mismatches 521; Indels 6; Gaps 2
XX
QY 32 ACATCAAGCAGGCGAATTTGACGCCCGGAGATTGAGATTGCCAGCAACATGCTCG 91
DB 355 ACATCGGCCTGGCTGCCCTGGGAGCGCAAGGCCCTGGACATTGCTGAGAACGAGATGCCGG 414
QY 92 CTCTGATTTCACTCAGGAACGCTGCTCAGGGGGAGAACCCCTTGCTGCTGCTAAATAG 151
DB 415 GCCTGATGCGTATGCGGAGCGGTACTCGGCTCCAAAGCCACTGAAGGGCGCCGCATCG 474
QY 152 TGGGCTGTACACACATCACAGCCAGACAGCGGTGTTGATTGAGACTCTGTGCCCTGG 211
DB 475 CTGGCTGCGCTGCACATGACCGGTGAGACGGCCGCTCCTCATTTGAGACCCCTGCACCCCTGG 534
QY 212 GGGCTCAGTGCCGCTGCTGCTGTTTAACATCTACTCAACTCAGATGAAGTAGCTGCAG 271
DB 535 GTGCTGAGGTGCAAGTGGTCCAGCTGCAACATCTTCTCCACCAGCAATGCGGGCGCTG 594
QY 272 CACTGGCTGAGGCTGAGTTGCACTGTTCCCTTGGAAAGGGCGAGTCAGAAAGATGACTTCT 331
DB 595 CCATTGCCAAGGCTGGCATTCGGGTGTATGCTGGAAGGGCGAAGCGAGAGAGTACC 654
QY 332 GGTGCTGATTTGACCGCTGTGTGAACATGATGGGTGGCAGGCCAATGATCCTGATG 391
DB 655 TGTGTGTCATTGAGCAGACCCCTGTACTTCAAGGACGAGGCCCTCAACATGATTCTGACG 714
QY 392 ATGGGGGAGACTTAACCCACTGGGTTTATAAGAGTATCCAAACGCTTTTAAGAAGATCC 451
DB 715 ACGGGGGGACCTCACCAACCTCATCCACACCAAGTACCCTGACGCTTCTGCCAGGCATCC 774
QY 452 GAGGCATTTGTGAAGAGAGCGGTACTGCTGTTCACAGGCTGTATCATCTCTCCAAAGCTG 511

```

Db	775	GAGGCACTCTGTAGAGAGACCACGACTGGGGTCCACAACTCTTACAAGATGATGGCCAAATG	834
Qy	512	GGAAGCTCTGTGTTCCGGCCATGAACGTCAATGATTCTGTACCACAAACAGAAATTGATA	571
Db	835	GGATCCTCAAGGTGCCCTGCCATCATGTCAATGACTCCGTACACCAAGACAAGTATTGACA	894
Qy	572	ACTTGTACTGTGCTGCCGAGAATCCATTTTGGATGGCCTGAAGAGACCACAGATGTGATGT	631
Db	895	ACCTCATATGGCTGCCGGAGTCCCTCATAGATGGCATCAAGCGGGCCACAGATGTGATGA	954
Qy	632	TTGGTGGAAACAAGTGGTGGTGTGTGGCTATGTTGAGAGGTAGGCAAGGGCTGTGTGCTG	691
Db	955	TTGCCGGCAAGGTAGCGGTGTGTAGCAGGCTATGTGTATGTGGCAAGGGCTGTGCCAGG	1014
Qy	692	CTCTCAAAAGCTCTTGGAGCAATTGTCTACATTACCGGAATGACCCCATCTGTGTCTGC	751
Db	1015	CCCTGGGGGGTTTCGGAGCCCGCTCATCATCACCGAGATTGACCCCATCAACGACACTGC	1074
Qy	752	AGGCTTCATGATGGGTTCAGGGTGGTAAAGCTAAATGAATCATCCGGCAAGTGCATG	811
Db	1075	AGGCTGCCATGGAGGGCTATGAGGTGACCAACCATGGATGAGGCCCTGCAGAGGGCAACA	1134
Qy	812	TCGTATACTTGCACAGGAATAGAATGTAGTGACACGGGACCACTTGGATCCGATGA	871
Db	1135	TCCTTGTACACCAACACAGGCTGTATTGTACATCATCTTGGCCGGCACTTTGAGCAGATGA	1194
Qy	872	AAAACAGTTGTATCGTATGCAATATGGGCCACTCCAACACAGAAATCGATGTGACCAGCC	931
Db	1195	AGGATGATGCCATTGTGTGTAACTTGACACTTTGACGTGAGATCGATGTCAAGTGGC	1254
Qy	932	TCCGCACTCCGAGCTGACGTGAGGAGCGAGTACGTTCTCAGGTGGAACCATGTCACTGGC	991
Db	1255	TCAACGAGAAACGCCGTGAGAGAGTGAACATCAAGCCGCGAGTGGACCGTATCGGTGA	1314
Qy	992	CAGATGGCAAAACGAGTTGTCTCTCCTGGCAGAGGGTCTGTACTCAATTTAGCTGCTCCA	1051
Db	1315	AGATGGGCGCCCGCATCATCTGTGTGGCCGAGGGTCCGGCTGTCAACCTGGGTGTGCCA	1374
Qy	1052	CAG---TTCCCACTTGTGTCTGTGCCATCACAGCCACACACAGGCTTTGGCACTGTATAG	1108
Db	1375	TGGGCCACCCCAAGCTTCGTGTATAGTAACTCTTCAACCAACGAGTGATGGCCAGATCG	1434
Qy	1109	AACCTTATAATGCAACCCGAGGGGCGATACAAGCAGGATGTACTTGTCTCTAAGAAAA	1168
Db	1435	AGCTGTGGACCCATCCAGACAAG---TACCCTGTTGGGGTTCATTCTGCCCCAAGAGCC	1491
Qy	1169	TGGATGAATACGTTGGCAGCTTGCATCTGCCATCATTTGATGCCACCTTACAGAGCTGA	1228
Db	1492	TGGATGAGCAGTGGCTGAAGCCCACTGGGCAAGCTGAATGTGAAGTTGACCAAGCTAA	1551
Qy	1229	CAGATGACCAAGCAAAATATCTGGGACTCAACAAAAATGGGCCATTCAAACCTAATTATT	1288
Db	1552	CTGAGAGCAAGCCCAAGTACCTGGGCAATGTCTCTGTGATGGCCCCCTTCAAGCCGATCACT	1611
Qy	1289	ACAGATAC 1296	
Db	1612	ACCGCTAC 1619	

Search completed: April 21, 2003, 00:05:52
Job time : 301.45 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2003, 23:48:04 ; Search time 57.932 Seconds
(without alignments)
6860.687 Million cell updates/sec

Title: US-09-782-051-1_COPY_549_1844

Perfect score: 1296
Sequence: 1 tccaaggcagcagcaattt.....aacctaattaccagatac 1296

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/lna/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/lna/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/lna/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/lna/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/lna/PTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/lna/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
C 1	406.4	31.4	636	4 US-09-328-111-850	Sequence 850, App
2	314.4	24.3	2211	4 US-09-318-448-26	Sequence 26, Appl
3	314.4	24.3	2211	4 US-09-347-878-2	Sequence 2, Appl
C 4	236.2	18.2	4411529	4 US-09-103-840A-1	Sequence 1, Appl
5	222.6	17.2	1812	2 US-08-669-536-1	Sequence 1, Appl
6	214.6	16.6	1767	3 US-08-930-894-1	Sequence 1, Appl
7	117	9.0	2226	2 US-08-896-005-2	Sequence 2, Appl
8	117	9.0	2226	4 US-09-347-878-3	Sequence 3, Appl
9	72.6	5.6	289	1 US-08-204-740-8	Sequence 8, Appl
10	72.6	5.6	289	3 US-09-081-167A-8	Sequence 8, Appl
11	72.6	5.6	289	3 US-09-081-395-8	Sequence 8, Appl
12	72.6	5.6	289	4 US-09-416-833-8	Sequence 8, Appl
13	72.6	5.6	289	5 PCT-US95-02521-8	Sequence 8, Appl
14	68.4	5.3	285	1 US-08-204-740-6	Sequence 6, Appl
15	68.4	5.3	285	3 US-09-081-167A-6	Sequence 6, Appl
16	68.4	5.3	285	3 US-09-081-395-6	Sequence 6, Appl
17	68.4	5.3	285	4 US-09-416-833-6	Sequence 6, Appl
18	68.4	5.3	285	5 PCT-US95-02521-6	Sequence 6, Appl
19	68.2	5.3	584	4 US-08-998-416-249	Sequence 249, App
C 20	41.8	3.2	7218	1 US-08-232-463-14	Sequence 14, Appl
21	40	3.1	50341	1 US-08-247-901C-1	Sequence 1, Appl
22	40	3.1	50341	2 US-09-075-904-1	Sequence 1, Appl
23	40	3.1	52297	4 US-09-426-436-1	Sequence 1, Appl
24	40	3.1	52297	4 US-08-705-557-1	Sequence 1, Appl
25	36	2.8	20137	4 US-09-262-773-206	Sequence 206, App
26	36	2.8	20137	4 US-09-262-773-9	Sequence 9, Appl
27	36	2.8	23071	4 US-09-262-773-210	Sequence 210, App

C 28	33.6	2.6	495	1 US-08-133-711-41	Sequence 41, Appl
C 29	33	2.5	50000	4 US-09-146-053-3	Sequence 3, Appl
30	32.8	2.5	31571	1 US-08-323-443B-1	Sequence 1, Appl
31	32.8	2.5	53526	3 US-08-658-136-2	Sequence 2, Appl
32	32.8	2.5	53577	3 US-08-658-136-1	Sequence 1, Appl
33	32.6	2.5	2101	1 US-08-106-761-1	Sequence 1, Appl
C 34	32.4	2.5	327	1 US-08-652-859-4	Sequence 4, Appl
C 35	32.4	2.5	327	1 US-08-919-706-4	Sequence 4, Appl
C 36	32.4	2.5	327	2 US-09-153-751-4	Sequence 6, Appl
37	31.4	2.4	4973	4 US-09-381-862-6	Sequence 108, App
38	31.2	2.4	233	2 US-08-687-080-108	Sequence 7, Appl
39	31	2.4	3306	4 US-09-770-170-7	Sequence 7, Appl
C 40	30.8	2.4	1614	4 US-09-522-217-7	Sequence 1, Appl
C 41	30.8	2.4	2665	3 US-09-040-005-1	Sequence 3, Appl
42	30.8	2.4	3186	3 US-08-863-102-3	Sequence 11, Appl
43	30.6	2.4	4049	1 US-08-162-809-17	Sequence 17, Appl
44	30.6	2.4	4097	1 US-08-162-809-11	Sequence 11, Appl
C 45	30.2	2.3	678	4 US-09-461-697-414	Sequence 414, App

ALIGNMENTS

RESULT 1
US-09-328-111-850/c
; Sequence 850, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dertl, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 850
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(636)
; OTHER INFORMATION: n - A,T,C or G
US-09-328-111-850

Query Match 31.4%; Score 406.4; DB 4; Length 636;
Best Local Similarity 99.5%; Pred. No. 2.5e-120;
Matches 418; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 159	TACACACATCACAGCCAGACAGC-GGTGTGATGAGACACTGTGCGCGGCGCTC	217
Db 420	TACACACATCACAGCCAGACAGCGGTGTGATGAGACACTGTGCGCGGCGCTC	361
QY 218	AGTCCCGCTGTCTGTGTAACATCTACTCACTCAGATGAAGTAGCTGCACACTGG	277
Db 360	AGTCCCGCTGTCTGTGTAACATCTACTCACTCAGATGAAGTAGCTGCACACTGG	301
QY 278	CTGAGGCTGAGGTGCACTGTTCGCTTGAAGGCGAGTCAGAAGTACTCTGTGCTGT	337

Db 300 CTGAGGCTGAGTTCAGTGTTCGCTTGGAAGGGGAGTCAAGATGACTTCTGCTGT 241
QY 338 GATATGACCGCTGTGTGAACATGGATGGGTGGCAGGCCAATGATCCTGATGATGGGG 397
Db 240 GATATGACCGCTGTGTGAACATGGATGGGTGGCAGGCCAATGATCCTGATGATGGGG 181
QY 398 GAGACTTAACCCACTGGGTTTATAGAAGTATCCAAAGTGTGTTAAGAAGATCCGAGCA 457
Db 180 GAGACTTAACCCACTGGGTTTATAGAAGTATCCAAAGTGTGTTAAGAAGATCCGAGCA 121
QY 458 TTGTGAAGAGAGCGTGTGTTTCAACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGC 517
Db 120 TTGTGAAGAGAGCGTGTGTTTCAACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGC 61
QY 518 TCTGTGTTCGGCCATGACGTCAATGATCTGTATACCAACAGAAAGTTGATTAATTGT 577
Db 60 TCTGTGTTCGGCCATGACGTCAATGATCTGTATACCAACAGAAAGTTGATTAATTGT 1

RESULT 2

US-09-318-448-26
; Sequence 26, Application US/09318448
; Patent No. 6210950
; GENERAL INFORMATION:
; APPLICANT: Johnson, William G.
; APPLICANT: Stenroos, Edward S.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; FILE REFERENCE: 601-1-057
; CURRENT APPLICATION NUMBER: US/09/318,448
; CURRENT FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; LENGTH: 2211
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-318-448-26

Query Match 24.3%; Score 314.4; DB 4; Length 2211;
Best Local Similarity 61.9%; Pred. No. 1.6e-90;
Matches 498; Conservative 0; Mismatches 306; Indels 0; Gaps 0;

QY 32 ACATCAAGCAGCAGAAATTGGACGCCGGAGATTGAGATTGACAGCAAGACATGCTG 91
Db 79 ACATCGGCTGCTGCTGCGGAGACGCAAGGCCCTGACATGCTGAGAACGATGCGG 138
QY 92 CTCTGATTTCACTCAGAAAGTGTCTCAGGGGAGAAAGCCCTGCTGCTAAATAG 151
Db 139 GCCTGATGCGTATGCGGAGCGGTACTCGGCCCTCCAAAGCCACTGAAGGGCCCGCATCG 198
QY 152 TGGGCTGTACACATCAGACGCCAGACAGCGGTGTGATGAGACACTCTGCTGCTGG 211
Db 199 CTGGCTGCTGACATGACCGTGGAGAGCGGCCCTCTCATTTGAGACCCCTGCTACCCCTGG 258
QY 212 GGGCTCAGTGGCGCTGCTGCTGTGTAACATCTACTCACTCAGATGAAGTAGCTGAG 271
Db 259 GTGCTGAGGTGAGTGTTCAGCTGCAACATCTTCTCCACCAAGAACATGCGGCGCTG 318
QY 272 CACTGGCTGAGCGTGAAGTGTGAGTGTTCGCTTGAAGGGCGAGTCAAGATGACTTCT 331
Db 319 CCATGGCCAAAGCGTGGCATTCGCGTGTATGCTGAAAGGGCGAAACGAGAGAGTACC 378
QY 332 GGTGCTGATGACCGCTGTGTGAACATGATGGGTGCAAGCCAAACATGATCTGATG 391
Db 379 TGTGTGATGATGAGCAGACCGCTGTACTCAAGGAGCGGCCCTCAACATGATCTGAGC 438
QY 392 ATGGGGGAGACTTAACCACTGGGTTTATAGAAGTATCCAAAGTGTGTTAAGAAGATCC 451
Db 439 ACGGGGGGAGACTTAACCACTGATCCACACCAAGTATCCCGCAAGCTTCTGCGGAGATCC 498
QY 452 GAGGCATTTGGAAGAGAGCGTGAAGTGTGCTCAGAGCTGTATCAGCTCTCCAAAGCTG 511
Db 511 GAGGCATTTGGAAGAGAGCGTGAAGTGTGCTCAGAGCTGTATCAGCTCTCCAAAGCTG 511

Db 499 GAGGCATCTCTGAGGAGACCAAGCACTGGGGGTCCACAACTCTACAGATGATGGCCAATG 558
QY 512 GGAAGCTCTGTGTTCGGCCATGAACGTCATATGATCTGTATCCAAACAGAGTTGATA 571
Db 559 GGAATCTCAAGGTGCGCTGCCATCAATGTCATATGATCTGCTCAACAGAGCAAGTTGACA 618
QY 572 ACTTGTACTGCTGCCGAGAAATCCATTTTGGATGAGGCTGGAAGAGCACAGATGTGATGT 631
Db 619 ACCTCTATGCTGCCGAGAGTCCCTCATATGATGATGATGATGATGATGATGATGATGATG 678
QY 632 TTGGTGGGAACAAGTGTGTGTGTGTGCTATGATGATGATGATGATGATGATGATGATGATG 691
Db 679 TTGGCGGCAAGGTAGCGGT 738
QY 692 CTCTCAAAAGCTCTGTGAGCAATGCTTACATTAACCGAAATGACCCCATCTGTGCTGTGC 751
Db 739 CCTTGGGGGGTTTGGAGCGCGCTCATCATCACCAGATTTGACCCCATCAACGCACTGTC 798
QY 752 AGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 811
Db 799 AGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 858
QY 812 TCGTAATAACTTGCACAGAAATA 835
Db 859 TCTTTGTACACACACAGCGCTGTA 882

RESULT 3

US-09-347-878-2
; Sequence 2, Application US/09347878C
; Patent No. 6376210
; GENERAL INFORMATION:
; APPLICANT: Yuan, Chong
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES
; FILE REFERENCE: 25885-1651
; CURRENT APPLICATION NUMBER: US/09/347,878C
; CURRENT FILING DATE: 1999-07-06
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2211
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Human S-adenosylhomocysteine hydrolase cDNA
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: M61831/GenBank
US-09-347-878-2

Query Match 24.3%; Score 314.4; DB 4; Length 2211;
Best Local Similarity 61.9%; Pred. No. 1.6e-90;
Matches 498; Conservative 0; Mismatches 306; Indels 0; Gaps 0;

QY 32 ACATCAAGCAGCAGAAATTGGACGCCGGAGATTGAGATTGACAGCAAGACATGCTG 91
Db 79 ACATCGGCTGCTGCTGCGGAGACGCAAGGCCCTGACATGCTGAGAACGATGCGG 138
QY 92 CTCTGATTTCACTCAGAAAGTGTCTCAGGGGAGAAAGCCCTGCTGCTAAATAG 151
Db 139 GCCTGATGCGTATGCGGAGCGGTACTCGGCCCTCCAAAGCCACTGAAGGGCCCGCATCG 198
QY 152 TGGGCTGTACACATCAGACGCCAGACAGCGGTGTGATGAGACACTCTGTGCGCTGG 211
Db 199 CTGGCTGCTGACATGACCGTGGAGAGCGGCCGTCTCTTGAAGACCTGCTACACCTGG 258
QY 212 GGGCTCAGTGGCGCTGCTGCTGTGTAACATCTACTCACTCAGATGAAGTAGCTGAG 271
Db 259 GTGCTGAGGTGAGTGTTCAGCTGCAACATCTTCTCCACCAAGAACATGCGGCGCTG 318
QY 272 CACTGGCTGAGGCTGAGTGTGAGTGTTCGCTTGAAGGGCGAGTCAAGAGATGACTTCT 331
Db 319 CCATTTGCCAAAGCGTGGCATTCGCGTGTATGCTGAAAGGGCGAAACGAGAGAGTACC 378

OY	332	GGTGGTGATTGACCGCTGTGTGAACAATGGATGGGTGGCAGGCCAACATGATCTGGATG	391
Db	379	TGTGTGCATTGAGCAGACACCCTGTACTTCAAAGGACGGGCCCTCAACATGATTTGGACG	438
OY	392	ATGGGGGAGACTTAACCCACTGGGTTTATAAGAAGTATCCAAACGTGTTAAGAAGATCC	451
Db	439	ACGGGGGCGACCTCACCAACCTCATCCACACCAAGTACCCGAGCTTCTGCCAGGCATCC	498
OY	452	GAGGCATTGTGGAAGAGAGCGGTGACTGTTGTTCACAGGCTGTATCAGCTCCCAAAGCTG	511
Db	499	GAGGCATCTCTGAGGAGACCACGACTGGGGGTCCACAACCTCTACAAAGATGATGCCCAATG	558
OY	512	GGAAGCTCTGTGTTCGCGGCATGAACGTCATGATTTCTGTTACCAACAGAAAGTTGATA	571
Db	559	GGATTCCTCAAGGTGCGCTGCCATCAATGTCATGACTCCGTCACCAAGAGCAAGTTGACA	618
OY	572	ACTTGTACTGCTGCGCAGAAATCCATTTGGATGGCTGGAAGAGCACACAGATGATGT	631
Db	619	ACCTCTATGGCTGCGGGAGTCCCTCATAGATGGCATCAAGCGGGCACAGATGTGATGA	678
OY	632	TTGGTGGGAACAAGTGTGTGTGTGGCTATGGTGAAGTAGGCAAGGGCTGTGTGCTG	691
Db	679	TTGGCGGCAAGGTAGCGGTGTAGCAGGGTATGTGATGTGGCAAGGGCTGTGCCACAGG	738
OY	692	CTCTCAAGCTCTTGAGCAATTTGTCTACATTAACGAATCGACCCCATCTGTCTCTGC	751
Db	739	CCCTGCGGGGTTTCGGAGCCCCGCTCATCATCACCGAGATTGACCCCATCAACGCACTGC	798
OY	752	AGGCCCTGCATGGATGGGTTCAAGGGTGTAAAGCTAAATGAAGTCATCCGGCAAGTCGATG	811
Db	799	AGGCTGCCATGGAAGGGCTATGAGGTGACCACCATGGATGAGGCCCTGTCAAGAGGGCAACA	858
OY	812	TCGTAAATACTTGACACAGGAATA	835
Db	859	TCCTTGTCAACCACACAGGCTGTA	882

```

RESULT 4
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

```

Query Match	Similarity	Score	DB	Length
Best Local	56.6%	Pred. No. 1.4e-63		
Matches	499	Conservative	0	Mismatches 373; Indels 9; Gaps 3;
QY	422	AGAGTATCCCAACGCTGTTTAAGAGAATCCGAGGCATTTGGAGAAGAGCGTGA	CTGGTG	481
Db	3629036	ACAAGTGGACCAAGATAGCCGAGTCGCTCAAGGGCGTCA	CCGAGAGACCACCACCGGCG	3628977
QY	482	TTTCACGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTGTTCGGCCATGAACGTCA		541
Db	3628976	TGCTGCGGCTCTACCAATTCGCCGCCGCCGGGATCTGGCCCTCCCGGCATCAACGTCA		3628917
QY	542	ATGATTTCTGTATCCCAACAGAAGTTGATACTGTACTGCTGCCGAGAAATCCATT	TTGG	601

Db 3628916 ACGACTCGGTGACCAAGTCCAAATTGACACAACAAGTACGGCACTCGGCACCTCCCTGATCG 3628857

QY 602 ATGGCTGAAGAGGACACAGATGTGATGTTGTGGGAAACAAGTGGTGTGTGCT 661

Db 3628856 ACGGCATCAACCGCGGCAACCGACGCGCTGATCGCGGGTAAGAAGTCTCATCTGCGCT 3628797

QY 662 ATGGTGAGGTAGGCAAGGCGCTGTGCTGTCTCAAGCCTGTGAGCAATTGTCTACA 721

Db 3628796 ACGGCGACGCTCGGTAAAGGCTGTGCGGAGGCGGATGAAGGCCAAGGAGCGCGGCTCCG 3628737

QY 722 TTACCGAATCGACCCCATCTGTGCTCTGCAAGGCGCTGCATGATGGGTTCAAGGTTGTA 781

Db 3628736 TCACCGAGATCGACCCCGATCAACGCGCTGCAGGCCCATGATGAGGCGCTTCGACGTGTCA 3628677

QY 782 AGCTAAATGAAGTCATCCGCGAAGTCGATGTCGTAATACCTGCACAGGAAATAAGATG 841

Db 3628676 CCGTCGAGGAGGCCCATCGGGACGCGGCACATCTGCTAAACCGCGACCGGCAACAAGACA 3628617

QY 842 TAGTGACACGGGAGCACCTTGGATCGCATGAAAAACAGTTGTATCGTATGCAATATGGCC 901

Db 3628616 TCATCATGCTCGAGCACAATTAAAGCGATGAAGACACACGCGATCTGGAAATATCGCC 3628557

QY 902 ACTCCAACACAGAAATCGATGACCAAGCCTCCGACCTCCGAGCTGACGTGGAGCGAG 961

Db 3628556 ACTTGACAACAGAGATCGACATGGCCGGGCTGAGCGCTCCGGGGCGACACGGGTCAACG 3628497

QY 962 TACGTTCTCAGGTGAGGACCATGTCATCTGGCCAGA--TGGCAACGAGTTGTCCTCTGG 1018

Db 3628496 TCAAGCCTCAGGTGACCTGTGSACTTTGGCGACACGGGCGCTCGATCATCTGTCTGT 3628437

QY 1019 CAGAGGGTCGTCTACTCAATTTAGCTGTCTCCACAGTTCCACC--TTGTCTGTCCA 1075

Db 3628436 CCGAGGGGCGGCTGCTGAACCTGGGCAATGCCACCGGGCACCCCTCTGCTGTATGAGCA 3628377

QY 1076 TCACAGCCACAACACAGAGGCTTGGCACTGATAGAACTCTATATGCAACCCGAGGGCGAT 1135

Db 3628376 ACAGCTTCGCTAACCAAGACGATCGCCCAAGATCGAGCTGTGACCAAGAACGACGAG--T 3628320

QY 1136 ACAAGCAGAGTGTGTACTTGTCTTCCTAAGAAAATGATGAATACGTGCCAGCTTGCAATC 1195

Db 3628319 ACGACAACGAGGTGTACCGGCTGCCCCAAGCACCTCGACGAGAGGTGGCTGGAATCCATG 3628260

QY 1196 TGCCATCATTTGATGCCCACTTAAGAGCTGACAGATGACCAAGCAAAATATCTGGAC 1255

Db 3628259 TCGAGGCCCTTGGCGGTCACTGACCAAGCTGACCAAGGAGAGGCCGAATACCTCGCG 3628200

QY 1256 TCACAAAAATGGGCCATTCAAACCTAATTATTACAGATAC 1296

Db 3628199 TCGACGTCGAGGGTCCCTTACAAGCCGGAACCACTACCGCTAC 3628159

RESULT 5
US-08-669-536-1
Sequence 1, Application US/08669536
Patent No. 5910444
GENERAL INFORMATION:
APPLICANT: MASUTA, CHIKARA
APPLICANT: UEHARA, KYOKO
APPLICANT: TANAKA, HIDEO
APPLICANT: KUMATA, SHIGERU
TITLE OF INVENTION: ORGANISMS IN WHICH THE EXPRESSION OF
TITLE OF INVENTION: S-ADENOSYLHOMOCYSTEINE HYDROLASE GENE IS INHIBITED
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
STREET: PO BOX 747
CITY: FALLS CHURCH
STATE: VA
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/669,536
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: MURPHY JR, GERALD M
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1254-128
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1812 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-08-669-536-1
```

Query Match 17.2%; Score 222.6; DB 2; Length 1812;
Best Local Similarity 56.4%; Pred. No. 3.7e-61;
Matches 499; Conservative 0; Mismatches 374; Indels 12; Gaps 4;

```
OY 424 AAGTATCCAAACGTTTGAAGATCCGAGCATGTGGAAGAGAGCGTGACTGTGTT 483
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 634 AATATATACCAAGATGAAGAAAGACTCGTGGTGTTCGAGGAAACTACCACTGAGTT 693
OY 484 CACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTTCCGGCCATGAACTGAT 543
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 694 AAGAGGCTTATCAGATGACGCTAATGGAATTTGCTTCCCTGCTATTATGTTAT 753
OY 544 GATTCTGTACCAACAGAGTTTGATTAATCTGTACTGTGCGCCGAGAAATCCATTGAT 603
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 754 GATTCTGTACCAAGAGAGTTGCAACAATGTACGAGTGCAGGCTGAGTGCCTGAT 813
OY 604 GGCCTGAGAGGAGCAGATGTGATGTTGGTGGGAAACAAGTGTGCTGTGCTAT 663
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 814 GGTCTCATGAGGCTACTGATGTTATGATGCGCGGAAAGCTTGCCTGTGCTGTTAT 873
OY 664 GTGAGGTAGGCAAGGCTGCTGTGCTGCTCTCAAGCTCTTGAGCAATTGCTACAT 723
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 874 GGAGATGTCGGCAAGGCTGTGCTGCTGCTGCTGAAACAGCCGCTGCTGATGCTG 933
OY 724 ACCGAATCGACCCCATCTGTGCTGCTGAGGCTGCAATGATGGTTCAGGGTGTAAAG 783
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 934 ACCGAGATTGACCTATCTGTGCTGCTGCTGAGGCTACCAATGGAAGCCCTCCAGTCTTACT 993
OY 784 CTAATGAGTCAATCCGCAAGTGCATGCTGTAATTAATCTGCACAGGAATAAGATGTA 843
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 994 CTAGAGGATGTCGTTCTGATGTTGATATCTTGTCAACACGACCGTACACAGACAT 1053
OY 844 GTGACACGGGAGCACTGGATCCGATGAAACAGTTGATCGTATGCAATATGGCCAC 903
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1054 ATCATGTTGACCAATGAGGAAGATGAAGAACATGCAATGTTTGCACATTTGCTCAC 1113
OY 904 TCCACACAGAAATCATGTGACCAAGCTCCGAC---TCCGAGCTGACGTTGGAGCGA 960
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1114 TTTGCAACGAAATCGATGCTTGTGCTGCAAGCTTACCTGCTGCTCAAGAGATGACA 1173
OY 961 GTACGTTCTCAGGTGACCATGTCATCTGGCCAGATGGCAACGAG---TTGCTCTCTG 1017
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1174 ATTAAGCTCAACCGACAGATGGGCTTCCCTGACACCAACAGTGCAATGCTCTTG 1233
OY 1018 GCAGAGGCTGCTCAATTTGAGCTGCTCCACAGTTCCACAC---TTTGTCTGTCC 1074
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1234 GCTGAGGCTGCTCAATTTGAGCTGCTCCACAGGACACCTAGTTTGTGATGCTG 1293
OY 1075 ATCACAGCCACACAGAGCTTTGGCACTGATAGACTCTATATGACAC---CGAGGGG 1131
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
DB 1294 TGCTGCTCACTAACCAAGTCAATGCCCACTCGAGTGTGGAATGAAGAAGACAGTGG 1353
OY 1132 CGATACAGCAGATGTGTACTGCTTCTTAAGAAATGATGATATACGTTGCCAGCTTG 1191
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1354 AAGTATGAGAAAGATGATGCTTCTTCCCAAAACACCTCGACGAGAGGTTGCTGACT 1413
OY 1192 CATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAACAAATATCTG 1251
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1414 CATCTCGAAAGCTCGGAGCCAAAGCTTACCAAACTTTGGAAGATCAAGCTGACTACAT 1473
OY 1252 GGACTCACAAAATGGCCATTCAAACCTAATTATTACAGATAC 1296
DB 1474 AGCGTCCAGTTGAGGGTCTTACAGCCTGCTCACTACAGGTAC 1518
```

RESULT 6
US-08-930-894-1
Sequence 1, Application US/08930894
Patent No. 6037524

GENERAL INFORMATION:

APPLICANT: GREENLAND, Andrew James

APPLICANT: DRAPER, John

APPLICANT: SKIPSEY, Marc

APPLICANT: WARNER, Simon

TITLE OF INVENTION: S-ADENOSYL-L-HOMOCYSTEIN HYDROLASE PROMOTER

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pillsbury Madison & Sutro

STREET: 1100 New York Avenue, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005-3918

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: MS Word

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/930,894

FILING DATE: 09-OCT-1997

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/GB96/00882

FILING DATE: 10-APR-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9507381.3

FILING DATE: 10-APR-1995

SEQUENCE CHARACTERISTICS:

LENGTH: 1767 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: SHH GENE FROM ASPARAGUS

FEATURE:

NAME/KEY: CDS

LOCATION: 26..1483

OTHER INFORMATION: /codon_start= 26

Query Match 16.6%; Score 214.6; DB 3; Length 1767;
Best Local Similarity 55.8%; Pred. No. 1.3e-58;
Matches 494; Conservative 0; Mismatches 379; Indels 12; Gaps 4;

```
OY 424 AAGTATCCAAACGTTTGAAGATCCGAGCATGTGGAAGAGAGCGTGACTGTGTT 483
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 596 AAGTACAGGAAGATGAAGATGAGATTGTCGGTGTGTCGAGAGACCAACCGGGGTC 655
```

QY 484 CACAGGCTGTATCAGCTCTCCAAAGCTGGAGCTCTGTGTTCCGGCCATGACGTCAAT 543
DB 656 AAGAGGCTTTACAGATGCAAGGCTAACAAATTCCTTTCTTCCCTGGATCAATGTCAAT 715
QY 544 GATTCTGTACCAACAGAGTTTGATTAATCTGCTGCTCCGAGAAATTCATTTGGAT 603
DB 716 GACTCCGTACCAAGCAAGTTTGACAATCTGTATGATGCGGACCTCTCTCCGAT 775
QY 604 GGCTGAAGAGGACAGATGTGATGTTGGTGGGAAACAAGTGTGTGTGTGCTAT 663
DB 776 GGCTGTATGAGGCGCACTGATGTATGATGCTGCAAGGTTGCAAGTGTGTGTGTAT 835
QY 664 GGTGAGTAGCAAGGCTGTGCTGTCTCTCAAGCTCTTGAGCAATGTCTACAT 723
DB 836 GGTGATGTGAGAGGCTGTGCTGTCTGCTGCAAGCAAGGCTGTGCTGTATGTG 895
QY 724 ACCGAATCGACCCATCTGTGCTGTGCAAGGCTGTGATGGGTTCAAGGTTGTAAG 783
DB 896 ACGGAGATCGACCCATCTGTGCTCTCAAGCCCTAATGAGGGCTCTCAAGGTTCTCACC 955
QY 784 CTAAATGAAGTATCCGCGCAAGTGTGCTGTCTAATTAATCTGACAGAAATAAGATGA 843
DB 956 CTCGAGATGTGTCTCAGAGGCGGATATCTTTGTAACCAACCGGTAACAAGGATC 1015
QY 844 GTGACACGGAGCACTTGATCGCATGAAAAACAGTTGATGATGCAATATGAGCCAC 903
DB 1016 ATCATGCTGACCAATGAGAGATGAGAACAATGCCATGTCTGCAACATTTGGCAC 1075
QY 904 TCCAACACAGAAATGATGTGACCAAGCTCCGAC--TCCGAGCTGACGTGGAGCCA 960
DB 1076 TTTGACACAGAGATGACATGCTAGTTTGAGACATACCCCTGGCATCAAGAGATCAC 1135
QY 961 GTACGTTCTCAGGTGACCATGTCTATGCGCCAGATGGCAA---ACGAGTTGCTCTCTG 1017
DB 1136 ATCAAGCCCCAGACTGACCGGTGGTCTTCCCTGAACCAACACTGTTAATTTGTTCTT 1195
QY 1018 GCAGAGGCTGCTTACTCAATTTGAGCTGCTCCACAGTT--CCACCTTTGTTGTTCC 1074
DB 1196 GCTGAGGCGGACTCATGAACCTTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1255
QY 1075 ATCAGAGCCACACAGAGCTTTGGCACTGATAGACTCTATATG--CACCCGAGGG 1131
DB 1256 TGCTCCTTACCAACAGGTGATGCTCAGCTAGAGTTGTGAATGAGAAGGCAAGCGGC 1315
QY 1132 CGATACAGCAGATGTGCTGCTCTAAGAAATGATGATACGTTGCCAGCTG 1191
DB 1316 AAGTATGAGAGAGGTTTACGTGCTCCCAAGCATCTTGATGAGAAGTAGCAGCGCTT 1375
QY 1192 CATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTG 1251
DB 1376 CACTTGGGCAAGCTGGAGCCCAAGCTTACAAAGCTCAGCCCTTACAGAGCGGACTACATC 1435
QY 1252 GGAATCAACAAATGGGCTTCAAACTTAATTAATACAGATAC 1296
DB 1436 AGCGTCCCATCGAGGTCCCTACAGCCACCTCACTACAGGTAC 1480

RESULT 7

US-08-896-005-2
Sequence 2, Application US/08896005
Patent No. 5854023
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Lal, Preeti
APPLICANT: Shah, Purni
TITLE OF INVENTION: S-ADENOSYL-L-HOMOCYSTEINE HYDROLASE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA

COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/896,005
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0337 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2226 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BLADTUT04
CLONE: 1519044
US-08-896-005-2

Query Match 9.0%; Score 117; DB 2; Length 2226;
Best Local Similarity 100.0%; Pred. No. 3e-27;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1180 GTTGCCAGCTTGCACTGCAATTTGATGCCACCTTACAGAGCTGACAGATGACCA 1239
DB 1 GTTGCCAGCTTGCACTGCAATTTGATGCCACCTTACAGAGCTGACAGATGACCA 60
QY 1240 GCAAAATATCTGGAGCTCAACAAATGGGCTTCAAACTTAATTAATACAGATAC 1296
DB 61 GCAAAATATCTGGAGCTCAACAAATGGGCTTCAAACTTAATTAATACAGATAC 117

RESULT 8

US-09-347-878-3
Sequence 3, Application US/09347878C
Patent No. 6376210
GENERAL INFORMATION:
APPLICANT: Yuan, Chong
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES
FILE REFERENCE: 25885-1651
CURRENT APPLICATION NUMBER: US/09/347,878C
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 2226
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(2226)
OTHER INFORMATION: Polynucleotide encoding human
OTHER INFORMATION: S-adenosyl-5-homocysteine hydrolase (SAHH) derived
OTHER INFORMATION: from bladder; n-a, c, g, or t
PUBLICATION INFORMATION:
PATENT DOCUMENT NUMBER: 08/896,005
PATENT FILING DATE: 1997-07-17
PUBLICATION DATE: 1998-12-29
US-09-347-878-3

Query Match 9.0%; Score 117; DB 4; Length 2226;

Best Local Similarity 100.0%; Pred. No. 3e-27;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1180 GTGCCAGCTTGATCTGCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAA 1239

Db 1 GTGCCAGCTTGATCTGCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAA 60

QY 1240 GCAAAATATCTGGACTCAACAAATGGCCATTCAAACCTAATTATTACAGATAC 1296

Db 61 GCAAAATATCTGGACTCAACAAATGGCCATTCAAACCTAATTATTACAGATAC 117

RESULT 9

US-08-204-740-8

; Sequence 8, Application US/08204740

; Patent No. 5753432

; GENERAL INFORMATION:

; APPLICANT: Gudkov, Andrei

; APPLICANT: Kazarov, Alexander

; APPLICANT: Mazo, Ilya

; APPLICANT: Roninson, Igor B

; TITLE OF INVENTION: Methods for Identifying Genetic

; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Alligretti & Witcoff, Ltd.

; STREET: 10 S. Wacker Drive, Suite 3000

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/204,740

; FILING DATE: 04-MAR-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: No. 5753432nan, Kevin E

; REGISTRATION NUMBER: 35,303

; REFERENCE/DOCKET NUMBER: 93,354-C

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 312-715-1000

; TELEFAX: 312-715-1234

; TELEX: 910-221-5317

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 289 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

US-08-204-740-8

Query Match 5.6%; Score 72.6; DB 1; Length 289;

Best Local Similarity 58.6%; Pred. No. 1.6e-13;

Matches 126; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 32 ACATCAAGCAGGAGAAATTGGACGCCCGGAGATTGAGATTGCAGAGCAACATGTCG 91

Db 75 ACATCGGCGCTGGCTGGCTGGGACGCAAGGCCCTGGAATTGCTGAGAACGAGATGCCG 134

QY 92 CTCTGATTCTACTCAGGAACGCTGCTCAGGGGAGAGCCCTTGCGTGGTCTAAATAG 151

Db 135 GCCTGATGCGTATGCGGAGCGGCTGCTGCTCCAAAGCCACTGAAGGGCGCCGATCG 194

QY 152 TGGCTGTACACATCAGAGCCAGACAGCGGTGTGATTGAGACACATCTGTCCTG 211

Db 195 CTGGCTGCTGACATGACCGGTGAGACGCGCCGCTCTCATTTGAGACCCCTGTCACCCCTG 254

QY 212 GGGCTCAGTCCCGCTGCTGCTGTTGATCATCTAC 246

Db 255 GTGCTGAGGTGACAGTGTGCTCAGCTGCAACATCTTC 289

RESULT 10

US-09-081-167A-8

; Sequence 8, Application US/09081167A

; Patent No. 6083745

; GENERAL INFORMATION:

; APPLICANT: Gudkov, Andrei

; APPLICANT: Kazarov, Alexander

; APPLICANT: Mazo, Ilya

; APPLICANT: Roninson, Igor B

; TITLE OF INVENTION: Methods for Identifying Genetic

; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant

; NUMBER OF SEQUENCES: 15

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff

; STREET: 300 S. Wacker Drive, 32nd Floor

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/081,167A

; FILING DATE: 18-MAY-1998

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: No. 6083745nan, Kevin E

; REGISTRATION NUMBER: 35,303

; REFERENCE/DOCKET NUMBER: 93,354-KK

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 312-913-0001

; TELEFAX: 312-913-0002

; TELEX:

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 289 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

US-09-081-167A-8

Query Match 5.6%; Score 72.6; DB 3; Length 289;

Best Local Similarity 58.6%; Pred. No. 1.6e-13;

Matches 126; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 32 ACATCAAGCAGGAGAAATTGGACGCCCGGAGATTGAGATTGCAGAGCAACATGTCG 91

Db 75 ACATCGGCGCTGGCTGGCTGGGACGCAAGGCCCTGGAATTGCTGAGAACGAGATGCCG 134

QY 92 CTCTGATTCTACTCAGGAACGCTGCTCAGGGGAGAGCCCTTGCGTGGTCTAAATAG 151

Db 135 GCCTGATGCGTATGCGGAGCGGCTGCTGCTCCAAAGCCACTGAAGGGCGCCGATCG 194

QY 152 TGGCTGTACACATCAGAGCCAGACAGCGGTGTGATTGAGACACATCTGTCCTG 211

Db 195 CTGGCTGCTGACATGACCGGTGAGACGCGCCGCTCTCATTTGAGACCCCTGTCACCCCTG 254

QY 212 GGGCTCAGTCCCGCTGCTGCTGTTGATCATCTAC 246

Db 255 GTGCTGAGGTGACAGTGTGCTCAGCTGCAACATCTTC 289

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02521
FILING DATE:

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:
LENGTH: 289 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
PCT-US95-02521-8

Query Match 5.6%; Score 72.6; DB 5; Length 289;
Best Local Similarity 58.6%; Pred. No. 1.6e-13;
Matches 126; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 32 ACATCAAGCAGCAGAAATTGGACGCCGAGATTGAGATTGACAGACAGACATGCTG 91
DB 75 ACATCGCGCTGGCTGCTGGGAGCGCAAGGCCCTGACATTGCTGAGAACGAGATGCCG 134
QY 92 CTCTGATTTCACTAGAAACGTCTCAGGGGGAGAGCCCTTGCTGCTGCTAAATAG 151
DB 135 GCCTGATGCGTATGCGGAGCGGTACTCGGCTCCAGCCACTGAAGGGCGCCGATCG 194
QY 152 TGGGCTGTACACACATCAAGACCCAGACAGCGGTGTGATTGAGACACTGTGCCCCG 211
DB 195 CTGGCTGCGCTGACACATGACCGTGGAGAGCGGCCGCTCTCATTTGAGACCCCTGTCACCCCTGG 254
QY 212 GGGCTCAGTGCCGCTGCTGCTGTACATCTAC 246
DB 255 GTGCTGAGGTGACAGTGTCTCCAGCTGCAACATCTTC 289

RESULT 14

US-08-204-740-6

Sequence 6, Application US/08204740

Patent No. 5753432

GENERAL INFORMATION:

APPLICANT: Gudkov, Andrei
APPLICANT: Kazarov, Alexander
APPLICANT: Mazo, Ilya
APPLICANT: Roninson, Igor B
TITLE OF INVENTION: Methods for Identifying Genetic
TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
TITLE OF INVENTION: Growth in Cancer Cells
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Allgrettl & Witcoff, Ltd.
STREET: 10 S. Wacker Drive, Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/204,740
FILING DATE: 04-MAR-1994
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: No. 5753432nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 93,354-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000

TELEFAX: 312-715-1234

TELEX: 910-221-5317

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
LENGTH: 285 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-204-740-6

Query Match 5.3%; Score 68.4; DB 1; Length 285;
Best Local Similarity 56.8%; Pred. No. 3.5e-12;
Matches 126; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 25 GTGAGAACATCAAGCAGCAGAAATTGGACGCCGAGATTGAGATTGACAGACAGAC 84
DB 64 GTGCGGACATCGACTGGCCCTGGGAGCGAAGGCTGTGATATAGCTGAGAAATGAG 123
QY 85 ATGCTGCTCTGATTTCACCTCAGAAACGTGCTCAGGGGAGAGACCCCTTGCTGTGCT 144
DB 124 ATGCCAGGCTTGATGCGCATGGGAGAGATGACTACAGCCCTCCAGCCACTGAAGGCTGT 183
QY 145 AAATAGTGGGCTGTACACACATCAAGCCAGACAGCGGTGTGATTGAGACACTGTGT 204
DB 184 CGCATTTGCTGCTGCTGCTGCGCATGACCGTGGAGACTGCTGTCTCATTTGAGACTCTCTG 243
QY 205 GCCCTGGGGCTCAGTGCCGCTGCTGCTGTGATCATCTAC 246
DB 244 GCCCTGGGTGCTGAGGGCGGCTGTCTCCAGCTGCAACATCTTC 285

RESULT 15

US-09-081-167A-6

Sequence 6, Application US/09081167A

Patent No. 6083745

GENERAL INFORMATION:

APPLICANT: Gudkov, Andrei
APPLICANT: Kazarov, Alexander
APPLICANT: Mazo, Ilya
APPLICANT: Roninson, Igor B
TITLE OF INVENTION: Methods for Identifying Genetic
TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
TITLE OF INVENTION: Growth in Cancer Cells
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
STREET: 300 S. Wacker Drive, 32nd Floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/081,167A
FILING DATE: 18-MAY-1998
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: No. 6083745nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 93,354-KK
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-0002
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 285 base pairs
TYPE: nucleic acid

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 02:39:45 ; Search time 259.419 seconds

(without alignments)
5018.577 Million cell updates/sec

Title: US-09-782-051-1_COPY_549_1844

Perfect score: 1296

Sequence: 1 tccaaggcgacgacattt.....aacctaattacagatac 1296

Scoring table: IDENTITY_NUC

Searched: 639749 seqs, 502280978 residues

Total number of hits satisfying chosen parameters: 1279498

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1296	100.0	2563	10 US-09-782-051-1	Sequence 1, Appl1
2	461.4	35.6	553	10 US-09-867-701-2023	Sequence 2023, Ap
3	406.4	31.4	636	10 US-09-879-536-850	Sequence 850, App
4	402.4	31.0	2200	10 US-09-925-301-217	Sequence 217, App
5	402.4	31.0	2429	12 US-10-044-090-344	Sequence 344, App
6	337	26.0	384	10 US-09-880-107-346	Sequence 346, App
7	241.6	18.6	1461	10 US-09-759-990-1	Sequence 1, Appl1
8	225.6	17.4	1422	9 US-09-738-626-836	Sequence 836, App
9	225.6	17.4	1557	9 US-09-746-650A-97	Sequence 97, Appl1
10	190	14.7	1877	12 US-10-021-121-1	Sequence 1, Appl1
11	186	14.4	1396	9 US-09-746-650A-101	Sequence 101, Appl1
12	174	13.4	720	9 US-09-738-626-838	Sequence 838, App
13	172.4	13.3	3830	9 US-10-037-598-26	Sequence 26, Appl1
14	172.4	13.3	513509	9 US-09-754-853A-4	Sequence 4, Appl1
15	168.2	13.0	433	10 US-09-960-352-4421	Sequence 4421, Ap
16	158	12.2	721	10 US-09-925-300-439	Sequence 439, App
17	150.8	11.6	382	10 US-09-960-352-11499	Sequence 11499, A
18	138.4	10.7	284	10 US-09-920-300A-1423	Sequence 1423, Ap
19	138.4	10.7	284	12 US-10-033-528-1423	Sequence 1423, Ap

20	137.6	10.6	424	10 US-09-960-352-9160	Sequence 9160, Ap
21	136.8	10.6	288	9 US-09-736-457-1179	Sequence 1179, Ap
22	136.8	10.6	288	9 US-09-902-941-1179	Sequence 1179, Ap
23	136.8	10.6	288	9 US-09-849-626-1179	Sequence 1179, Ap
24	136.8	10.6	288	9 US-10-017-754-1179	Sequence 1179, Ap
25	136.4	10.5	288	10 US-09-998-598-2191	Sequence 2191, Ap
26	128.2	9.9	352	10 US-09-960-352-12928	Sequence 12928, A
27	124.4	9.6	283	10 US-09-960-352-7006	Sequence 7006, Ap
28	114.4	8.8	793	10 US-09-966-881-9	Sequence 9, Appl1
29	111.4	8.6	708	9 US-09-738-626-837	Sequence 837, Appl1
30	108.2	8.3	389	10 US-09-960-352-10073	Sequence 10073, A
31	102.4	7.9	728	10 US-09-770-149-117	Sequence 117, App
32	100.6	7.8	237	10 US-09-960-352-14145	Sequence 14145, A
33	98.4	7.6	263	10 US-09-923-876-477	Sequence 477, App
34	96	7.4	410	10 US-09-960-352-1365	Sequence 1365, App
35	94.2	7.3	253	10 US-09-960-352-7284	Sequence 7284, App
36	91.4	7.1	399	10 US-09-878-574-838	Sequence 838, App
37	81.4	6.3	219	10 US-09-960-352-12043	Sequence 12043, A
38	79.2	6.1	456	10 US-09-960-352-2551	Sequence 2551, App
39	78.2	6.0	835	10 US-09-770-445-679	Sequence 679, App
40	75.2	5.8	405	10 US-09-960-352-5161	Sequence 5161, App
41	72.6	5.6	289	10 US-09-799-946-8	Sequence 8, Appl1
42	69.8	5.4	347	10 US-09-960-352-2035	Sequence 2035, App
43	68.4	5.3	285	10 US-09-799-946-6	Sequence 6, Appl1
44	67.6	5.2	163	10 US-09-878-574-9524	Sequence 9524, Ap
45	65.8	5.1	428	10 US-09-960-352-10555	Sequence 10555, A

ALIGNMENTS

RESULT 1
US-09-782-051-1
Sequence 1, Application US/09782051
Patent No. US20020035078A1
GENERAL INFORMATION:
APPLICANT: Hart, Derek N J
TITLE OF INVENTION: Enzyme having S-adenosyl-L-homocysteine hydrolase
TITLE OF INVENTION: (AHCY) type activity
FILE REFERENCE: 24305 MRB
CURRENT APPLICATION NUMBER: US/09/782,051
CURRENT FILING DATE: 2001-02-14
PRIOR APPLICATION NUMBER: PCT/NZ97/00133
PRIOR FILING DATE: 1997-10-06
PRIOR APPLICATION NUMBER: NZ 299507
PRIOR FILING DATE: 1996-10-04
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 2563
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (3)..(1847)
OTHER INFORMATION: Open reading frame extends without a stop codon
OTHER INFORMATION: for the full 5' nucleotide sequence. The
OTHER INFORMATION: Initiation codon has yet to be identified.
US-09-782-051-1

Query Match 100.0%; Score 1296; DB 10; Length 2563;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	TCCAAGGCGACGACCAATTCTGTGTGAAGACATCAAGCAGGAGAAATTTGACGCCG 60
DB	549	TCCAAGGCGACGACCAATTCTGTGTGAAGACATCAAGCAGGAGAAATTTGACGCCG 608
QY	61	GAGATTGAGATTGACAGACATGTCTGTGATTTCACCTCAGGAACGCTCAG 120
DB	609	GAGATTGAGATTGACAGACATGTCTGTGATTTCACCTCAGGAACGCTCAG 668
QY	121	GGGAGAGAGCCCTTGGCTGTCTAATAATAGTGGGCTGTACACATCACAGCCAGACA 180

Db 669 GGGAGAGAGCCCTGGCTGGTCTAAATAGTGGGCTGTACACACATCACAGCCACAGACA 728
QY 181 GCGGTGTGATGAGACACTCTGTGCCCTGGGGGCTCAGTGGCGCTGTCTGTGTAC 240
Db 729 GCGGTGTGATGAGACACTCTGTGCCCTGGGGGCTCAGTGGCGCTGTCTGTGTAC 788
QY 241 ATCTACTCACTCAGATGAAGTAGCTGCAGCACTGGCTGAGGCTGAGTGCAGTGTTC 300
Db 789 ATCTACTCACTCAGATGAAGTAGCTGCAGCACTGGCTGAGGCTGAGTGCAGTGTTC 848
QY 301 GCTTGAAGGCGGAGTCAGAGATGACTTCTGTGTGTATGTACCCGCTGTGTACATG 360
Db 849 GCTTGAAGGCGGAGTCAGAGATGACTTCTGTGTGTATGTACCCGCTGTGTACATG 908
QY 361 GATGGGTGGCAGGCCAATCATGATCCTGGATGATGGGGAGACTTAACCCACTGGGTTAT 420
Db 909 GATGGGTGGCAGGCCAATCATGATCCTGGATGATGGGGAGACTTAACCCACTGGGTTAT 968
QY 421 AAGAATATCCAAAGCTGTTAAGAGATCCGAGCATGTGGAAGAGAGAGCTGATGCT 480
Db 969 AAGAATATCCAAAGCTGTTAAGAGATCCGAGCATGTGGAAGAGAGAGCTGATGCT 1028
QY 481 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAGCTGTGTCCGGCATGAACGTC 540
Db 1029 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAGCTGTGTCCGGCATGAACGTC 1088
QY 541 AATGATTTCTGTACCAACAAGATTGATTAATCTGCTGCTGCCAGAAATTCATTTTG 600
Db 1089 AATGATTTCTGTACCAACAAGATTGATTAATCTGCTGCTGCCAGAAATTCATTTTG 1148
QY 601 GATGGCTGAAGAGAGACACAGATGTGATGTGGTGGGAAACAAGTGTGTGTGGC 660
Db 1149 GATGGCTGAAGAGAGACACAGATGTGATGTGGTGGGAAACAAGTGTGTGTGGC 1208
QY 661 TATGTTGAGTAGGCAAGGGCTGTGTGTCTCAAAAGCTCTTGAGCAATTTGTCTAC 720
Db 1209 TATGTTGAGTAGGCAAGGGCTGTGTGTCTCAAAAGCTCTTGAGCAATTTGTCTAC 1268
QY 721 ATTACCGAAATCGACCCCATCTGTGTCTGCGAGGCTGCATGATGGGTTCAGGGTGTGTA 780
Db 1269 ATTACCGAAATCGACCCCATCTGTGTCTGCGAGGCTGCATGATGGGTTCAGGGTGTGTA 1328
QY 781 AAGCTAAATGAAGTATCCGGCAAGTGCATGTCTAATACTTGACAGGAAATGAAGAT 840
Db 1329 AAGCTAAATGAAGTATCCGGCAAGTGCATGTCTAATACTTGACAGGAAATGAAGAT 1388
QY 841 GTAGTGACACGGGAGCACTTGATCGCATGAATAACAGTTGTATCGTATGCAATATGGGC 900
Db 1389 GTAGTGACACGGGAGCACTTGATCGCATGAATAACAGTTGTATCGTATGCAATATGGGC 1448
QY 901 CACTCCAACACAGAAATCGATGTGACCAAGCCTCCGCACTCCGAGCTGACGTGGAGCGA 960
Db 1449 CACTCCAACACAGAAATCGATGTGACCAAGCCTCCGCACTCCGAGCTGACGTGGAGCGA 1508
QY 961 GTACGTTCTCAGGTGAGCATGTCTATCTGGCCAGATGGCAAAAGTGTCTCTCTGGCA 1020
Db 1509 GTACGTTCTCAGGTGAGCATGTCTATCTGGCCAGATGGCAAAAGTGTCTCTCTGGCA 1568
QY 1021 GAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTTCCACCTTTGTCTGTCTCACA 1080
Db 1569 GAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTTCCACCTTTGTCTGTCTCACA 1628
QY 1081 GCCACAACACAGGCTTTGGCACTGATAGAACTCTAATAATGACCCGAGGGGCGATACAAG 1140
Db 1629 GCCACAACACAGGCTTTGGCACTGATAGAACTCTAATAATGACCCGAGGGGCGATACAAG 1688
QY 1141 CAGGATGTGACTTGTCTCTAAGAAATGATGAATACGTTGCCAGCTTGCACTGCGCA 1200
Db 1689 CAGGATGTGACTTGTCTCTAAGAAATGATGAATACGTTGCCAGCTTGCACTGCGCA 1748
QY 1201 TCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGAGTCAAC 1260

Db 1749 TCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGAGTCAAC 1808
QY 1261 AAAATGGCCATTCAAACCTAATATATACAGATAC 1296
Db 1809 AAAATGGCCATTCAAACCTAATATATACAGATAC 1844

RESULT 2

US-09-867-701-2023
; Sequence 2023, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aylate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; NUMBER OF SEQ ID NOS: 2001-05-29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2023
; LENGTH: 553
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(553)
; OTHER INFORMATION: n = A,T,C or G
US-09-867-701-2023

Query Match 35.6%; Score 461.4; DB 10; Length 553;
Best Local Similarity 97.3%; Pred. No. 8e-147; Matches 510; Conservative 0; Mismatches 9; Indels 5; Gaps 4;

QY 774 GGTGTTAAGCTAAATGAAGTATCCGGCAAGTGCATGCTAATACTTGACAGGAAA 833
Db 1 GGTGTTAAGCTAAATGAAGTATCCGGCAAGTGCATGCTAATACTTGACAGGAAA 60
QY 834 TAAGATGTAGTACACAGGAGCACTTGATCGCATGAATAACAGTTGTATGTCGAA 893
Db 61 TAAGATGTAGTACACAGGAGCACTTGATCGCATGAATAACAGTTGTATGTCGAA 120
QY 894 TATGGGCACTCCAAACAGAAATCGATGTGACCAAGCCTCCGCACTCCGAGCTGAGTG 953
Db 121 TATGGGCACTCCAAACAGAAATCGATGTGACCAAGCCTCCGCACTCCGAGCTGAGTG 180
QY 954 GAGCGAGTACGTTCTCAGGTGAGCATGTCTATCTGGCCAGATGGCAACGAGTGTCT 1013
Db 181 GAGCGAGTACGTTCTCAGGTGAGCATGTCTATCTGGCCAGAT -GMAAACGAGTGTCT 239
QY 1014 CCTGGCAGAGGCTGCTACTCAATTTGAGCTGCTCCACAGTTCCACCTTTGTCTGTC 1073
Db 240 CCTGGCAGAGGCTGCTACTCAATTTGAGCTGCTCCACAGTTCCACCTTTGTCTGTC 299
QY 1074 CATCACAGCCACACAGGCTTTGGCACTGATAGAACTCTAATAATGACCCGAGGGCG 1133
Db 300 CATCACAGCCACACAGGCTTTGGCACTGATAGAACTCTAATAATGACCCGAGGGCG 359
QY 1134 ATACAAGCAGATGTGACTGCTCTCTAAGAAA -TGATGAATACGTTGCCAGCTGTC 1192
Db 360 ATACAAGCAGATGTGACTGCTCTCTAAGAAAATGATGAATACGTTGCCAGCTGTC 419
QY 1193 ATCTGCCATCATTTGATGCCCACTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1252
Db 420 ATCTGCCATCATTTGATGCCCACTTACAGAGCTGACAGATGACCAAGC -AAATATCTGG 478
QY 1253 GACTCAACAAAATGGCCATTCAAACCTAATATATACAGATAC 1296
Db 479 GACTCAACAAA -TGGGCAATTCACCACTAATATATACAGATAC 520

```
RESULT 3
US-09-879-536-850/c
; Sequence 850, Application US/09879536
; Patent No. US20020144298A1
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dertl, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/879,536
; PRIOR APPLICATION NUMBER: 2001-09-21
; PRIOR FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 850
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(636)
; OTHER INFORMATION: n - A,T,C or G
US-09-879-536-850

Query Match          31.4%; Score 406.4; DB 10; Length 636;
Best Local Similarity 99.5%; Pred. No. 5.2e-128;
Matches 418; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 159 TACACACATCACAGCCCAAGACG-GGTGTGATGAGACACTCTGTGCCCTGGGGCTC 217
DB 420 TACACACATCACAGCCCAAGACGCGGGTGTGATGAGACACTCTGTGGGGCTC 361
QY 218 AGTGGCGGTGTGTGTGTAACATCTACTCACTCAGAAATGAAGTAGCTGCAGACTGG 277
DB 360 AGTGGCGGTGTGTGTGTAACATCTACTCACTCAGAAATGAAGTAGCTGCAGACTGG 301
QY 278 CTGAGGCTGAGTGTGAGTGTGCTTGGAGGGCGAGTCAGAAAGATGACTTCTGTGGT 337
DB 300 CTGAGGCTGAGTGTGAGTGTGCTTGGAGGGCGAGTCAGAAAGATGACTTCTGTGGT 241
QY 338 GTATTGACCGCTGTGTGAACATGATGGGTGGCAGGCCAATGATCTGTGATGGGG 397
DB 240 GTATTGACCGCTGTGTGAACATGATGGGTGGCAGGCCAATGATCTGTGATGGGG 181
QY 398 GAGACTTAACCCACTGGGTTTATAAGAGTATCCAAACGTTTAAAGAGATCCGAGCA 457
DB 180 GAGACTTAACCCACTGGGTTTATAAGAGTATCCAAACGTTTAAAGAGATCCGAGCA 121
QY 458 TTGGAAGAGAGAGCGTGTGTTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGC 517
DB 120 TTGGAAGAGAGAGCGTGTGTTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGC 61
QY 518 TCTGTGTTCGGCCATGAACGTATGATCTGTATACCAAGAGAGTTGATACTGTGT 577
DB 60 TCTGTGTTCGGCCATGAACGTATGATCTGTATACCAAGAGAGTTGATACTGTGT 1

RESULT 4
US-09-925-301-217
; Sequence 217, Application US/09925301
; Patent No. US20020052308A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 217
; LENGTH: 2200
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2188)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-217

Query Match          31.0%; Score 402.4; DB 10; Length 2200;
Best Local Similarity 58.4%; Pred. No. 2.6e-126;
Matches 741; Conservative 0; Mismatches 521; Indels 6; Gaps 2;

QY 32 ACATCAAGCAGGAGATTTGAGCGCGGAGATTTGAGATTTGAGACAGACATGCTCTG 91
DB 108 ACATCGCGCTGGCTGCGTGGGAGCCAGCGCCCTGGACATTTGCTGAGAACGAGATGCCG 167
QY 92 CTCTGATTTCACTCAGGAACGCTGCTCAGGGGAGAGGCCCTTGGCTGTCTAAATAG 151
DB 168 GCGTATGCGTATGCGGAGCGGCTACTCGCCCTCCAAAGCCACTGAAGGGCGCCGCTGC 227
QY 152 TGGGCTGTACACACATCACAGCCAGACAGCGGTGTGATTTGAGACACTCTGTGCCCTGG 211
DB 228 CTGGCTGCTGACATGACCGGTGAGACGCGCCCTCCTCATTTGAGACCCCTGTCACCCCTGG 287
QY 212 GGGCTCAGTCCGCTGTGTGTGTAACATCTACTCACTCAGAAATGAAGTAGCTGCAG 271
DB 288 GTGCTGAGTGTGAGTGTGTCAGCTGCAACATCTCTCCACCCAGACCATGCGCGGCTG 347
QY 272 CACTGGCTGAGGCTGAGTGTGAGTGTGCTTGGAGGGCGGAGTCAGAGATGACTTCT 331
DB 348 CCATTTGCAAGGCTGGCATTTCCGGTGTATGCTGGAAGGGCGCAACGAGAGAGTACC 407
QY 332 GGTGTGTATTGACCGCTGTGTGTAACATGATGGGTGGCAGGCCAATGATCTGTGATG 391
DB 408 TGTGTGATTGAGCAGACCCCTGTACTTCAAGGAGCGGGCCCTCAACATGATTTCTGAGC 467
QY 392 ATGGGAGACTTAACCCACTGGGTTTATAAGAGTATCCAAACGTTTAAAGAGATCC 451
DB 468 ACGGGGCGACCTCACCAACCTCATCCACCAAGTACCCGCGCTTCTGCGAGCATCC 527
QY 452 GAGGCTTTGGAAGAGAGCGTGTGTTTCACAGGCTGTATCAGCTCTCCAAAGCTG 511
DB 528 GAGGCTTTGGAAGAGAGCGTGTGTTTCACAGGCTGTATCAGCTCTCCAAAGCTG 587
QY 512 GGAAGCTCTGTGTTCCGGCCATGAACGTCATGATTTCTGTTACCAACAGAGTTGATA 571
DB 588 GGATCTTCAAGGTGCGCTGCCATCATGTCAATGACTCCGCTCACCACAGAGATTGACA 647
QY 572 ACTTGTACTGCTGCCGAGATCCATTTTGGATGGCTGGAAGAGACCAAGATGTGATGT 631
DB 648 ACCCTTATGGCTGCCGGAGTCCCTCATAGATGGCATCAAGCGGGCCACAGATGTGATGA 707
QY 632 TTGGTGGAAACAAGTGTGTGTGTGCTATGTTGAGGTGAGGCAAGGGCTGTGTGCTG 691
DB 708 TTGCGGCAAGGTAGCGGTGTGAGAGGCTATGTTGATGTGGCAAGGGCTGTGCCAGG 767
QY 692 CTCTCAAGCTCTTGGAGCAATGCTTACATTTACCGAATGACCCCATCTGTGCTGCTGC 751
DB 768 CCTGTGGGGGTTTGGAGCCCGCGTCATCATCACCAGAGATTGACCCCATCAACGCACTGC 827
```

QY	752	AGGCTGCATGATGGGTTCAAGGTGGTAAAGCTAAATGAAGTCATCCGGCAAGTCATG	811
Db	828	AGGCTGCATGAGGGCTATGAGGTGACCACCATGGATGAGGCTGTCAAGAGGGCAACA	887
QY	812	TCGTATAACTTGACAGGAATAAGATGTAGTACACGGGAGCATTGGATCGCATGA	871
Db	888	TCTTGTGCACCACCAAGGCTGTATTGACATCATCTTGGCCGGCAGCTTGAAGAGATGA	947
QY	872	AAACAGTTGTATCGTATGCAATATGGGCCACTCCACACAGAATCGATGTGACCAGCC	931
Db	948	AGGATGATGCCATTGTGTGTACATTGGACACCTTTGACCTGGAGATCGATGTCAAGTGGC	1007
QY	932	TCCGCACTCCGAGCTGACGTGGGAGCGAGTACGTTCTCAGGTGGACCATGTCACTTGGC	991
Db	1008	TCAACGAGAACGCCGTGGAGAAAGGTGAACATCAAGCCGCCAGGTGGACCGGTATCGGTTGA	1067
QY	992	CAGATGGCAACGAGTTGTCTCTCTGGCAGAGGGTCTACTCAATTTGAGCTGCTCCA	1051
Db	1068	AGATGGGCGCGCATCATCTGTCTGGCCGAGGGTGGCTGTCAACCTGGGTTGTGCCA	1127
QY	1052	CAG--TTCCACACCTTTGTCTGTCCATCACAGCCACAACAAGGCTTTGGCACTGATAG	1108
Db	1128	TGGGCCACCCACAGCTTCGTGATGAGTAACCTCTTACCACAACGAGTGATGGCGCAGATCG	1187
QY	1109	AACTCTATAATGCACCCCGAGGGCGGATACAAAGCAGATGTGTACTTGCCTTCCTAAGAAA	1168
Db	1188	AGCTGTGACCCATCCAGACAAG--TACCCCGTTGGGGTTTCAATTTCTGCCCCAAGAAGC	1244
QY	1169	TGGATGAATACGTTGCCAGCTTGCACTGCCATCATTTGATGCCACCTTACAGAGCTGA	1228
Db	1245	TGGATGAGGCAGTGGCTGAAGCCCACTGGGCAAGCTGAATGTGAAGTTGACCAAGCTAA	1304
QY	1229	CAGATGACCAAGCAAAATATGTGGGACTCAACAAAAATGGGCCATTCAAACTAATTAAT	1288
Db	1305	CTGAGAAGCAAGCCCAAGTACTGGGGCATGTCTGTGATGGCCCTTCAAGCCGATCACT	1364
QY	1289	ACAGATAC 1296	
Db	1365	ACCGCTAC 1372	

RESULT 5
US-10-044-090-344
; Sequence 344, Application US/10044090

```

? APPLICANT: Olga Bandman
? TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
? FILE REFERENCE: PA-0028 US
? CURRENT APPLICATION NUMBER: US/10/044, 090
? CURRENT FILING DATE: 2002-01-09
? NUMBER OF SEQ ID NOS: 850
? SOFTWARE: PERL Program
? SEQ ID NO 344
? LENGTH: 2429
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc_feature
? OTHER INFORMATION: Incyte ID NO. US20020137081A1 1468237CBI
US-10-044-090-344

```

Query Match	31.0%;	Score 402.4;	DB 12;	Length 2429;
Best Local Similarity	58.48;	Pred. No. 2.8e-126;		
Matches 741; Conservative	0;	Mismatches 521;	Indels 6;	Gaps 2.

QY 32 ACATCAAGCAGCCAGAATTGGACGCCCGGAGATTGAGATTGCAGAGCAAGACATGTCTG 91
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 108 ACATCGGCGCTGGCTGCTGGGGACGCAAGGCCCTGGACATTGCTGAGAACGAGATGCCGG 167
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 92 CTCTGATTTCACTCAGGAACGTGCTCAGGGGGGAGACCCTTGCGTGTAAATAAG 151
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db	168	GCCTGATCCGTTATCGCGGAGCGGTACTCGGCTCCAAAGCCACTGAAGGGCGCCGCATCG	227
QY	152	TGGGCTGTACACATCACAGCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGG	211
Db	228	CTGGCTGCTGCACATGACCGGTGAGACGCGCTCTCTCATTTGAGACCCTCGTCACCCTGG	287
QY	212	GGGCTCAGTGCCTGCTGCTCTGCTTGTAACTCTACTCACTCAGAATGAAGTAGCTGCAG	271
Db	288	GTCCTGAGGTGCAGTGTCTCCAGCTGCACACATCTTCTCCACCAGGACCATTGCGGGGCTG	347
QY	272	CACGTGGCTGAGGCTGGAGTTGCAGTGTCTGCTTGGGAGGGCGAGTCAGAAGATGACTTCT	331
Db	348	CCATTGCGCAAGGCTGGCATTTCCGGTGTATGCCCTGGAAGGGCGAAGCGGAGGAGTACC	407
QY	332	GGTGTGTATTGACCGCTGTGTGAACATGGATGGGTGGCAGGCCAACATGATCTGTGATG	391
Db	408	TGTGTGATGAGCAGAACCTGTACTTCAAGGACGGGCGCCCTCAACATGATTTCTGGACG	467
QY	392	ATGGGAGACTTAACCCACTGGGTTTATAGAAGTATCCAAAGCTGTTTAAGAAGATCC	451
Db	468	ACGGGGCGACCTCAACCACTCATCCACACCAAGTACC CGCAGCTTCTGCCAGCATCC	527
QY	452	GAGCATTTGTGAAGAGAGCGTGACTGTGTTCACAGGCTGTATCAGCTCTCCAAAGCTG	511
Db	528	GAGGCATCTCTGAGGAGAGCACCAGCACTGGGGTCCACACACCTCTACAGATGATGCCAATG	587
QY	512	GGAAGCTCTGTGTCCGGCCATGAACGTCATGATTTCTGTACCAACAGAAAGTTTGATA	571
Db	588	GGATCCTCAAGGTGCTTCGCATCAATGTCAATGACTCCGTCACCAAGCAAGTTTGACA	647
QY	572	ACTGTACTGCTGCCGAGATCCATTTGGATGGCCTGAAGAGGACACAGATGTGATGT	631
Db	648	ACCTTATGCTGCCGGAGTCCCTCATAGATGGCATCAAGCGGGCCACAGATGTGATGA	707
QY	632	TTGTGGAAACAAGTGTGTGTGTGGCTATGTGTAGGTAGGCAAGGCTGCTGTGCTG	691
Db	708	TTGCCGGCAAGTAGCGGTGTGAGCAGGCTATGTGATGTGGCAAGGCTGTGCCAGG	767
QY	692	CTCTCAAGCTCTTGAGACAATTGTCTACATTACCAGAAATCGACCCCATCTGTGCTGCG	751
Db	768	CCCTGCGGGTTTCGGAGGCCCGCTCATCATCACCAGATTGACCCCATCAACGCACCTGC	827
QY	752	AGGCTGCATGATGGGTTCAGGGTGTAAAGCTAAATGAAGTCATCCGGCAAGTCGATG	811
Db	828	AGGCTGCCATGAGGGCTATGAGGTGACCACCATGATGAGGCTGTACAGAGGGCAACA	887
QY	812	TCGTATACTTGACAGAAATAAGATGTAGTGACACGGGAGCACTTGCATCGCATGA	871
Db	888	TCTTTGTCAACACACAGGCTGTATTGACATCATCTTGGCCGGCACCTTGTGACAGATGA	947
QY	872	AAACAGTGTATCGTATGCAATATGGGCCACTCCACACAGAAATCGATGTGACCAGCC	931
Db	948	AGGATGATGCCATTGTGTGAACATTGGACACTTTGACGTGGAGATCGATGTCAAGTGCC	1007
QY	932	TCCGCACTCCGAGCTGACGTGGAGCGAGTAGCTTCTCAGGTGGACCATGTCTGCGC	991
Db	1008	TCAACGAGAAGCCGTGAGAAAGGTGAACATCAAGCCGAGGTGAGACCGGTATCGGTGA	1067
QY	992	CAGATGGCAACAGAGTTGCTCTCTGCGCAGAGGGTCTACTCAATTGAGCTGCTCCA	1051
Db	1068	AGAATGGGGCGGCATCATCTGCTGGCGGAGGGTCCGCTGTCAACCTGGGTGTGCCA	1127
QY	1052	CAG--TTCCACACCTTTGTCTGTCTCATTCACAGCCACACACAGGCTTTGGCACTGATAG	1108
Db	1128	TGGGCCACCCACAGCTTGTGATGAGTACTCTTCAACCAACAGGTGATGGCGCAGATCG	1187
QY	1109	AACTCTAATATGACCCCGAGGGCGATACAGCAGGATGTGTACTTGTCTTCTTAAGAAA	1168
Db	1188	AGCTGTGAGACCATCCAGACAAG--TACCCCGTTGGGGTTTCATTTCTGCCCCAAGAAC	1244
QY	1169	TGGATGAATACGTGGCAGCTTGCATCTGCCATCATTTGATGCCACCTTACAGAGCTGA	1228
Db	1245	TGGATGAGGCAGTGGCTGAAGCCCACTTGGCCAAAGCTGAATGTGAAGTTGACCAAGCTAA	1304

OY	1229	CAGATGACCAAGCAAATAATCTGGGACTCACAACAAAATGGGCCATTCAAAACCCTAATTATT	1288
Dd	1305	CTGAGAAGCAAGCCAGTACTGTGGCATGTCTGTGATGGCCCCCTTCGAAGCCGGATCACT	1364
OY	1289	ACAGATAC	1296
Dd	1365	ACCGCTAC	1372

RESULT 6

```

US-09-880-107-346
; Sequence 346, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 346
; LENGTH: 384
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA157401
US-09-880-107-346

```

Query Match	26.08;	Score 337;	DB 10;	Length 384;
Best Local Similarity	99.08;	Pred. No. 1.9e-104;		
Matches 381; Conservative	0;	Mismatches 0;	Indels 4;	Gaps 4;

QY	444	GAAGATCCGAGGCATTTGTGGAAGAGAGCGGTGACTGGTGTTCACAGGCTGTATCAGCTCTC	503
Db	1	GAAGATCCGAGGCATTTGTGGAAGAGAGCGGTGACTGGTGTTCACAGGCTGTATCAGCTCTC	60
QY	504	CAAAAGCTGGGAAGCTCTGTGTCCGGCCATGAACGTCATGATTCGTACCAACAGAA	563
Db	61	CAAAAGCTGGGA - CTCGTGTTCGGCCATGAACGTCATGATTCGTACCAACAGAA	119
QY	564	GTTTGATAACTTGTACTGCTGCCGAGAATCCATTTTGGATGGCCCTGAAGAGACCACAGA	623
Db	120	GTTTGATAACTTGTACTGCTGCCGAGAATCCATTTTGGATGGCCCTGAAGAGACCACAGA	179
QY	624	TGTGATGTTTGGTGGGAACAAGTGTGTGTGGCTATGGTGAAGTAGGCAAGGGCTG	683
Db	180	TGTGATGTTTGGTGGGAACAAGTGTGTGTGGCTATGGTGAAGTAGGCAAGGGCTG	239
QY	684	-CTGTGCTGCTCTC - AAAGCTCTTGAGCAATTGTCTACATTACCGAAATCGACCCCATC	741
Db	240	CCTGTGCTGCTCTCAAAAGCTCTTGAGCAATTGTCTACATTACCGAAATCGACCCCATC	299
QY	742	TGT ³² CTCTGACAGGCCCTGCATGATGGGTTCAAGGGTGT - AAAGCTAAATGAAGTCAATCCG	800
Db	300	TGTGCTCTGCAGGCCCTGCATGATGGGTTCAAGGGTGTAAAGCTAAATGAAGTCAATCCG	359
QY	801	GCAAGTCGATGTCGTAATAACTTGC	825
Db	360	GCAAGTCGATGTCGTAATAACTTGC	384

RESULT 7
US-09-759-990-1
; Sequence 1, Application US/09759990

```

; Patent No. US20020119491A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Anticancer, Inc.
;
; APPLICANT: Xu, Mingxu
;
; APPLICANT: Han, Qinghong
;
; TITLE OF INVENTION: HIGH EXPRESSION AND PRODUCTION OF HIGH
; TITLE OF INVENTION: SPECIFIC ACTIVITY RECOMBINANT S-ADENOSYLHOMOCYSTEINASE
; TITLE OF INVENTION: (SAHH) AND IMPROVED ASSAYS FOR S-ADENOSYLMETHIONINE (SAM)
; FILE REFERENCE: 31276-20026.00
; CURRENT APPLICATION NUMBER: US/09/759,990
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/176,444
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
;
; SEQ ID NO 1
;
; LENGTH: 1461
;
; TYPE: DNA
;
; ORGANISM: Unknown
;
; FEATURE:
;
; OTHER INFORMATION: Nucleotide sequence encoding SAHH
;
; US-09-759-990-1

```

Query Match	18.6%;	Score 241.6;	DB 10;	Length 1461;
Best Local Similarity	57.8%;	Pred. No. 1.6e-71;		
Matches 490; Conservative	0;	Mismatches 349;	Indels 9;	Gaps 3;

QY	454	GGCATTGTGGAAGAGAGCGCTGACTGCTGTTCACAGCGCTGTATCAGCTCTCCAAAGCTGGG	513
Db	613	GGTGTTCGGAAGAGACAACAACAGGTGTCCACCGCCTCTACCAAGCTCGAAGAGGGC	672
QY	514	AAGCTCTGTGTTCCGGCCATGAACGTCAATGATGTCTGTACCAACAAGAAGTTGATTAAC	573
Db	673	AAACTCCTCTTCCACGACCATCAACGTCAACGACGCGTGTACAAGTCCAAGTTCGATTAAC	732
QY	574	TTGTACTGCTGCCGAGAAATCCATTTTGATGGCCTGAAGAGACACAGATGTGATGTTT	633
Db	733	ATCTACGGCTGCCCGCACTCCCTTATCGATGGTATCAACCGTGCTTCCGATGTCAATGATC	792
QY	634	GGTGGAAACAAGTGTGTGTGTGCTATGTTGAGGTAGGCAAGGCTGTGTGCTGCT	693
Db	793	GGCGGCAAGACAGCTCTCTCATGGGTTACGGCGATGTGGCAAGGCTGCGCTCAATCC	852
QY	694	CTCAAAAGCTCTTGAGACAATTGTCTACATTAACGAAATGACCCCATCTGTGCTGTGACG	753
Db	853	CTCCGTGGCCAGAGCGCTCCGCTATATCATCACAGAACTCGAACCCATCTGCGCTCTCCAG	912
QY	754	GCCTCATGGATGGGTTCAGGGGTGTTAAAGCTAAATGAATCATCCGGCAAGTCGATGTC	813
Db	913	GCTGCCATGGAAGGCTACAGGTCCGCCGATCGAGGAAGTCGTCGAAGGATGTGATATC	972
QY	814	GTAATACTTGACACAGAAATAAGAATGTAGTACACGGAGACCACTTGATCGCATGAATA	873
Db	973	TTCGTTACATGCACAGAAACTGCGATATCATCTCTGTGACATGATGGCCAGATGAAG	1032
QY	874	AACAGTTGATCGTATGCATATATGGGCCACTCCACACAGAAATCGATGTGACCAAGCTC	933
Db	1033	GATAAGGCTATTGTTCGGTAACATCGGCCACTTCGATAACGAAATTGATACAGATGGCCTC	1092
QY	934	CGCA--CTCCGGAGCTGACGTGGGAGCGAGTAGCTTCAGGTGGACCATGTCAATCGG	990
Db	1093	ATGAATAATCCAGGCATCAACACATCCCAATCAAGCCAGAAATACGACATGTGGAAATTC	1152
QY	991	CCAGATGGCAAAAGAGTTGTCCTCTCGCAGAGGGGTGCTACTCAATTTGAGCTGTCC	1050
Db	1153	CCAGATGGCCACGCTATCTCTCTTCTGTGAGGGCGGCTTCTTAACCTTGGCTGGGCT	1212
QY	1051	ACAGTT--CCCACCTTGTGTCTGTCCATCACAGCCACAACAAGGCTTTGGCACTGATA	1107
Db	1213	ACAGGTCACCCATCTTTCGTTATGTCAATGTCAATTCACAAACAGACACTCGCTCAGCTC	1272
QY	1108	GAACCTCTAATATGCACCCGAGGGGCGATACAAAGCAGAGATGTGTAATCTTGCCTCTTAAGAA	1167

Db 1273 GACCTCTACGA---AAAGAGAGGAATCTCGAGAAGAGGTTTACACACTTCCGAAGCAT 1329
QY 1168 ATGGATGAATACGTTGCCAGCTTGCAATCTGCCATCATTTGATGCCCCACCTTACAGAGCTG 1227
Db 1330 CTCGATGAAGAAGTCGCTCGCCCTCCACCTCGGATCTCTCGATGTCACCTTTACAAGCTT 1389
QY 1228 ACAGATGACCAAGCAAAATATCTGGGACTCAACAAAATGGCCATTCAAACTAATTAT 1287
Db 1390 ACACAGAAGCAGGCTGACTACATCAACGTTCCAGTTGAGGGTCTTACAAAGTCTGATGCT 1449
QY 1288 TACAGATA 1295
Db 1450 TACCGTTA 1457

RESULT 8
US-09-738-626-836

; Sequence 836, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738, 626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 836
; LENGTH: 1422
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-836

Query Match 17.4%; Score 225.6; DB 9; Length 1422;
Best Local Similarity 56.8%; Pred. No. 4.6e-66;
Matches 477; Conservative 0; Mismatches 354; Indels 9; Gaps 3;

QY 463 GAAGAGAGCGTGAAGTGTGTCACAGGCTGTATCACTCTCCAAAGCTGGAGCTCTGT 522
Db 586 GAGGAACACACACCGGTGTGACCGCTGTACCACTTCGCTGAAGAAGCGCTGCTGCT 645
QY 523 GTCCGGCATGAAGTCAATGATTCTGTACCAACAGAGTTTGATACTTGTACTGC 582
Db 646 TTCCACGCGATGAAGTCAAGAGCGTGTACCAAGTCAAGTTGATAACAAGTACGGC 705
QY 583 TGCCGAGATCCATTTTGGATGGCCCTGAAGAGACACAGATGTATGTTGGTGGAAA 642
Db 706 ACCCGCCACTCCCTGATGACGCGCATCAACCGCCGCTGACATGCTCATGGCGGCAAG 765
QY 643 CAAGTGTGTGTGTGCTATGTGAGGTAGGCAAGGCTGCTGCTCTCAAGCT 702
Db 766 AACGTGCTGTGTGCGGTACGGCGATGTGCGCAAGGGCTGCGCTGAGGCTTTGACGGC 825
QY 703 CTGGAGCAATTGTCTACATTAACGAATGACCCCATCTGCTCTGACAGGCTGATG 762
Db 826 CAGGGCGCTGCGCTCAAGGTCAACGAAGCTGACCCCAATCAACGCTCTTCAGGCTGATG 885
QY 763 GATGGCTTCAGGGTGTAAAGCTAAATGAAGTCAATCCGGCAAGTGCATGTCTAATAACT 822

Db 886 GATGGCTACTCTGTGTGTCACCGGTGTGATGAGGCCATCGAGAGCGCCGACATCGTATCAC 945
QY 823 TGCACAGGAATTAAGATGTAGTGACACGGGAGCACTTGATGCGATGAACAGTGT 882
Db 946 GCGACCGCAACAGACATCATTTCTTCTGAGAGATGCTCAAGATGAAGATCAACGCT 1005
QY 883 ATCGTATGCAATATGGCCACTCCACACACAGAAATCGATGTG---ACCAGCTCCGCACT 939
Db 1006 CTGCTGGGCAACATCGGTCACTTTGATATGAGATCGATATGCATTCCTGTGTGACCGC 1065
QY 940 CCGAGCTGACGTGGGAGCGAGTACGTTCTCAGGTGACCATGTCTATCTGCGCATGCG 999
Db 1066 GACGACGTACCCCGACACGATCAAGCCACAGGTGACAGAGTCAACCTTCTCCACCGGT 1125
QY 1000 AAACGAGTGTCTCTCTGGCAGAGAGGGTCTCTACTCAATTGAGCTGCTCCAC---AGTT 1056
Db 1126 CGCTCATCATCGTCTCTGTCGGAAGGTGCGCTGTTGAACCTTGCAACGCGCACCGAGAC 1185
QY 1057 CCCACCTTGTCTGTGTCATCACAGCCACACACAGGCTTTGGCACTGATAGACTCTAT 1116
Db 1186 CCATCATTTGTCTATGTCACACTCTTTCGCGATCAGACCATTTGCGCAGATCGAATGTC 1245
QY 1117 AATGACCCGAGGGCGATACAAAGAGATGTGTACTTCTCTAAGAAATGATGAA 1176
Db 1246 CA---AAACGAAGACAGATACGAGAGAGGTCTACCGTCTGCTAAGGTCTCGACGAA 1302
QY 1177 TACGTGCGAGCTTGCATCTGCCATCATTTGATGCCCCACCTTACAGAGCTGACAGATGAC 1236
Db 1303 AAGGTGGCAGCATCCAGCTTGAGGCTCTCGGGGCTGACGCTCAACGACCAAGAGAG 1362
QY 1237 CAAGCAAAATATCTGGGACTCAACAAAATGGGCCATTCAAACCTAATTATTACAGATAC 1296
Db 1363 CAGGCTGAGTACATCGGCTTGACGTTGACAGGCCCATTCAGCCGAGACACTACCGCTAC 1422

RESULT 9
US-09-746-660A-97

; Sequence 97, Application US/09746660A
; Publication No. US20030049804A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberer, Gregor
; APPLICANT: Kim, Jun-Won
; APPLICANT: Lee, Heung-Schick
; APPLICANT: Hwang, Byung-Joon
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
; FILE REFERENCE: BGI-121CP2
; CURRENT APPLICATION NUMBER: US/09/745, 660A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 09/606740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 09/603124
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142101
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/148613
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/187970
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19931420.9
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn Vers. 2.0
; SEQ ID NO 97
; LENGTH: 1557
; TYPE: DNA

```

; ORGANISM: Corynebacterium glutamicum
;
; FEATURE:
;
; NAME/KEY: CDS
;
; LOCATION: (101)..(1534)
;
; OTHER INFORMATION: RXN00132
;
US-09-746-660A-97

```

Query Match	17.4%;	Score 225.6;	DB 9;	length 1557;
Best Local Similarity.	56.8%;	Pred. No. 4.9e-66;		
Matches 477; Conservative	0;	Mismatches 354;	Indels 9;	Gaps 3

QY	463	GAAGAGAGCGGTGACTGGTGTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGT	522
Db	698	GAGGAACCAACCACCGGTGTGCACCCGCTGTACCACCTTCGCTGAAGAAGGCGTGCCT	757
QY	523	GTTCCGGCCATGAACGTCATGATTTCTGTATCCAAACAGAAGTTGATAACTGTACTGC	582
Db	758	TTCCAGCGATGAACGTCACAGACGCTGTCAACCAAGTCCAGTTGATAACCAAGTACGGC	817
QY	583	TGCCGAGATCCATTTGGATGGCCTGAAGAGAGACCAGATGTGATGTTGGTGGAAA	642
Db	818	ACCCGCCACTCCCTGATCGACGGCATCAACCGCGCACTGACATGCTCATGGCGGCAAG	877
QY	643	CAAGTGTGTGTGTGGCTATGGTGAAGTAGGCAAGGCGCTGTGTGCTCTCAAGCT	702
Db	878	AACGTGCTGTCTGGCGTTACGGCGATGTGGCAAGGCGTGGCTGAGGCTTTCGACGGC	937
QY	703	CTTGAGCATTTGTCATATTACCGAATCGACCCCATCTGTGCTCTCGAGGCGCTGCATG	762
Db	938	CAGGGCGCTCGCGTCAAGGTCACCGAAGCTGACCCCATCAACGCTCTTCAGGCTCTGATG	997
QY	763	GATGGTTCAGGGTGGTAAGCTAAATGAAGTCATCCGCGCAAGTCGATGTCGTAATACT	822
Db	998	GATGGCTACTCTGTGTGTCAACCGTTGATGAGGCCATCGAGACGCCGACATCGTATCACC	1057
QY	823	TGCACAGGAATAAGATGTAGTGACACGGGACACTTGGATCGCATGAAAAACAGTTGT	882
Db	1058	GCGACCGCAACAGACATCATTTCTTCGAGCAGATGCTCAAGATGAAGATCACGCT	1117
QY	883	ATCGTATGCAATATGGGCCACTCCAACACAGAAATCGATGTG--ACGAGCCTCCGCACT	939
Db	1118	CTGCTGGCAACATCGGTCACTTTGATAATGAGATGATATGCATTCCTGTTGCACCGC	1177
QY	940	CCGAGCTGACGTGGGAGCGAGTACGTTCTCAGGTGACCATGTCACTGCGCAGATGGC	999
Db	1178	GACGACGTCAACCCGCAACGATCAAGCCACAGGTGCGAGAGTTCACCTTCTCCACCGGT	1237
QY	1000	AAACGAGTGTCTCTCTGGCAGAGGGGTGCTCACTCAATTTGAAGCTGCTCCAC--AGTT	1056
Db	1238	CGCTCCATCATCGTCTCTGTCCGAAGTGCCTGTTGAACCTTGGCAACGCCACCGGACAC	1297
QY	1057	CCCACCTTGTGTCTGTCCATCACAGCCACACACAGGCTTTGGCACTGATAGAATCTAT	1116
Db	1298	CCATCATTTGTCAATGCCAACTCTTTGCCGATCAGACCAATGCGCAGATCGAACTGTTC	1357
QY	1117	AATGACCCGAGGGCGGATACAAGCAGAGATGTACTGCTTCCCTAAGAAATGATGAA	1176
Db	1358	CA--AAACGAAGGACAGTACGAGAAGAGGTCTACCGCTGTGCTTAAGGTCTCGACGAA	1414
QY	1177	TACGTTGCAGCTTGCATCTGCCATCATTTGATGCCCACTTACAGAGCTGACAGATGAC	1236
Db	1415	AAGGTGGACGCAATCGACGTTGAGGCTCTCGGGGTCAAGTCAACCGAATGACCAAGGAG	1474
QY	1237	CAAGCAAAATATCTGGGACTCAACAAAAATGGGCCATTCAACTAATTATTACAGATAC	1296
Db	1475	CAGGCTGATACATCGGCGTTGACGTTGCAGGCCCCATTAAGCCGGAACACTACCGCTAC	1534

RESULT 10
US-10-021-121-1/c
; Sequence 1, Application US/10021121
; Patent No. US2002014244A1
; GENERAL INFORMATION:

APPLICANT: Caras, Ingrid W
TITLE OF INVENTION: A2-1 Neurotrophic Factor
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/021,121

FILING DATE: 06-Dec-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/635,130

FILING DATE: 19-Mar-1996

ATTORNEY/AGENT INFORMATION:

NAME: Torchla, Phd., Timothy E.

REGISTRATION NUMBER: 3

REFERENCE/DOCKET NUMBER: P1001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/225-8674

TELEFAX: 650/952-9881

FORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1877 base pairs

TYPE: Nucleic Acid

STRANDEDNESS: Double

TOPOLOGY: Linear

FEATURE:

NAME/KEY: Extra Cellular Domain

LOCATION: 244-895

IDENTIFICATION METHOD:

OTHER INFORMATION:

FEATURES:

NAME/KEY: Transmembrane Domain

NAME/ AEL: LAUSILLO
LOCATION: 001-078

IDENTIFICATION METHOD:
LOCATION: 901-9/8

IDENTIFICATION METHOD

OLH

FEATURE:

NAME/KEY: signal

LOCATION: 244-321

IDENTIFICATION MET

OTHER INFORMATION:

US-10-021-121-1

Query Match	14.7%;	Score 190;	DB 12;	Length 1877;
Best Local Similarity	100.0%;	Pred..No. 7.9e-54;		
Matches 190; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	735	CCCCATCTGTGCTCTG	CAGGCCCTGCATG	ATGGGTTCAAGGGTG	TAAAGCTAAATGA	AGT	794
Db	1446	CCCCATCTGTGCTCTG	CAGGCCCTGCATG	ATGGGTTCAAGGGTG	TAAAGCTAAATGA	AGT	1387
QY	795	CATCCGGCAGTGCATG	TCGTAATAACTTGC	CACAGGAAATAAGA	ATGTAGTGACACGG	GGA	854
Db	1386	CATCCGGCAGTGCATG	TCGTAATAACTTGC	CACAGGAAATAAGA	ATGTAGTGACACGG	GGA	1327
QY	855	GCACTTGGATCGCATG	AAACAGTTGTATCG	TATGCAATATG	GGCCACTCCAA	CACAGA	914
Db	1326	GCACTTGGATCGCATG	AAACAGTTGTATCG	TATGCAATATG	GGCCACTCCAA	CACAGA	1267
QY	915	AATCGATGTG	924				
Db	1266	AATCGATGTG	1257				

RESULT 11

US-09-746-660A-101
; Sequence 101, Application US/09746660A
; Publication No. US20030049804A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; APPLICANT: Kim, Jun-Mon
; APPLICANT: Lee, Heung-Schick
; APPLICANT: Hwang, Byung-Joon
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
; TITLE OF INVENTION: METABOLIC PATHWAY PROTEINS
; FILE REFERENCE: BGI-121CP2
; CURRENT APPLICATION NUMBER: US/09/746,660A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 09/606740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 09/603124
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142101
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/148613
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/187970
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn Vers. 2.0
; SEQ ID NO:101
; LENGTH: 1396
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(1396)
; OTHER INFORMATION: FRXA01371
US-09-746-660A-101

Query Match 14.4%; Score 186; DB 9; Length 1396;
Best Local Similarity 57.3%; Pred. No. 1.5e-52;
Matches 377; Conservative 0; Mismatches 275; Indels 6; Gaps 2;

QY 463 GAAGAGCGTGAAGTGTTCACAGCGTGTATCAGCTCTCCAAAGCTGGAGCTCTGT 522
DB 698 GAGGAAACACACACCGGTGTGACCGCTGTACCACTCGTGAAGAAGCGTGCCT 757
QY 523 GTCCGGCCATGAAGTCAATGATTTCTGTACCAAGAAGTTGATACTGTACTGC 582
DB 758 TTCCAGCGATGAAGTCAAGACGCTGTACCAAGTTCAGTTGATAACAAGTACGGC 817
QY 583 TGCCGAGATCCATTTTGGATGGCCTGAAGAGGACCAAGATGTGATTTGGTGGAAA 642
DB 818 ACCCGCCACTCCCTGATCAAGCGCATCAACCGCCACTGACATGCTCATGGCGGCAAG 877
QY 643 CAAGT 702
DB 878 AACGTGCTGT 937
QY 703 CTGAGCAATTTGTACATTACCGAATCGACCCCATCTGTCTGTGAGGCGCTGCATG 762
DB 938 CAGGGCGCTCGCTCAAGGTCAAGCAAGCTGACCAATCAACGCTCTTCAAGGCTGTG 997
QY 763 GATGGCTTCAAGGTGTAAAGCTAAATGAAGTCAAGCGCAAGTGCATGTGTAATAACT 822
DB 998 GATGGCTACTCTGT 1057

QY 823 TGCACAGGAATAGATGTAGTACACAGGAGCACTTGGATCGCATGAAAAACGTTGT 882
DB 1058 GCGACCGGCAACAGGACATCATTTCTTCGAGCAGATGCTCAAGATGAAGATCAGCT 1117
QY 883 ATCGTATGCAATATGGGCACTCCACACACAGAATGATGTGACCAAGCTC--CGCACT 939
DB 1118 CTGCTGGCAACATCGTCACTTTGATATGAGATGATGATTCCTGTTGCAAGC 1177
QY 940 CCGAGCTGACGTGGAGCGAGTACGTTCTCAGGTGACCATGTCATCTGGCAGATGCC 999
DB 1178 GACGAGCTACCCCGCACACAGATCAAGCCACAGGTGACAGATTCACCTCTCCACGGT 1237
QY 1000 AAACGAGTGTCTCTCGGACAGAGGCTGCTACTCAATTTGAGCTGCTCCACAGTTCC 1059
DB 1238 CGCTCCATCATCTGCTGTCGGAAGGTGCGCTGTGAACTTGGCAAGCCGACGAC 1297
QY 1060 ACC--TTGTCTGTCTCATACAGCCACACACAGGCTTTGGCACTGATAGACTCT 1114
DB 1298 CCATCATTTGTCTCATGTCACACTCTTTCGCCGATCAGACCATTTGCCAGATCGACTGT 1355

RESULT 12

US-09-738-626-838/c
; Sequence 838, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 838
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-838

Query Match 13.4%; Score 174; DB 9; Length 720;
Best Local Similarity 61.0%; Pred. No. 1.3e-48;
Matches 282; Conservative 0; Mismatches 180; Indels 0; Gaps 0;

QY 463 GAAGAGCGTGAAGTGTTCACAGCGTGTATCAGCTCTCCAAAGCTGGAGCTCTGT 522
DB 465 GAGGAAACACACACCGGTGTGACCGCTGTACCACTCGTGAAGAAGCGTGCCT 406
QY 523 GTCCGGCCATGAAGTCAATGATTTCTGTACCAAGAAGTTGATACTGTACTGC 582
DB 405 TTCCAGCGATGAAGTCAAGACGCTGTACCAAGTTCAGTTGATAACAAGTACGGC 346
QY 583 TGCCGAGATCCATTTTGGATGGCCTGAAGAGGACCAAGATGTGATTTGGTGGAAA 642
DB 345 ACCCGCCACTCCCTGATCAAGCGCATCAACCGCCACTGACATGCTCATGGCGGCAAG 286
QY 643 CAAGT 702
DB 285 AACGTGCTGT 226

QY	703	CTTGGAGCAATTGCTCTACATTAACGGAAATCGACCCCATCTGTGCTCTGCAAGGCCCTGCATG	762
Db	225	CAGGGCGCTCGCGTCAAGGTCACCGAAGCTGACCCCAATCAACGCTCTTCAGGCTCTGATG	166
QY	763	GATGGGTCAGGGTGTGTAAGCTAAATGAAGTCATCCGGCAAGTCGATGCTGTAATAACT	822
Db	165	GATGGCTACTCTGTGTCACCGTGTGATGAGGCCATCGAGGACGCCGACATCGTGATCACCC	106
QY	823	TGCACAGGAATAAGAAATGTAGTGAACACGGGAGCACCTTGGATCCGATGAATAACAAGTTGT	882
Db	105	GCGACCCGGCAACAAGGACATCATTTCTTCGAGCAGATGCTCAAGATGAAGGATCAACGCT	46
QY	883	ATCGTATGCAATATGGGCCACTCCAACACAGAAATCGATGTG	924
Db	45	CTGCTGGGCAACATCGGTCACTTTGATTAATGAGATCGATATG	4

RESULT 13

```

; Sequence 26, Application US/10037598
; Patent No. US20020157143A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto CO
; APPLICANT: Concibido, Vergel
; APPLICANT: Delanney, Xavier
; TITLE OF INVENTION: Soybean plants with Enhanced yields and Methods for Breeding for
; TITLE OF INVENTION: Screening of Soybean plants with Enhanced yields
; FILE REFERENCE: 38-21(52175)B
; CURRENT APPLICATION NUMBER: US/10/037,598
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 06/260,040
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
; LENGTH: 3830
; TYPE: DNA
; ORGANISM: Glycine max
US-10-037-598-26

```

Query Match	13.38;	Score 172.4;	DB 9;	Length 3830;
Best Local Similarity	55.28;	Pred. No. 1.2e-47;		
Matches 423; Conservative	0;	Mismatches 331;	Indels 12;	Gaps 4;

QY	543	TGATCTGTACCAACAGAGTTTGATACTTGACTGCTCCGAGAAATCCATTTTGA	602
Db	3062	TTAATCTGTGTTTGTAAACAGTTTGACAACCTGTATGGGTGCCGTCACTCTCCCTGA	3121
QY	603	TGGCCTGAAGAGGACACAGATGTGATGTTTGGTGGAAACAAGTGGTGTGTGGCTA	662
Db	3122	TGCTCTCATGAGGGCTACCGATGTTATGATTTGCTGGAAAGGTGGCTGTTGTGGCTGATTA	3181
QY	663	TGCTGAGGTAGGCAAGGCTGCTGTGCTGCTCTCAAAGCTCTTGAGCAATTGCTTACAT	722
Db	3182	TGGTGATGTTGGCAAGGTTGTGCTGCTGCAATGAAGCAGGCTGGTGTGCTGTCTATCGT	3241
QY	723	TACCGAATCGACCCCATCTGTGCTCTGCAGGCTGCATGGATGGGTTCAAGGTGTGTA	782
Db	3242	GACCGAGATTGATCCCATCTGTGCCCCCTTCAGGCTTCATGGAAGCCCTTCAGGTTCTGAC	3301
QY	783	GCTTAATGAAGTCATCCGGCAAGTCGATGTCGTAATTACTTGCACAGGAATAAGATGT	842
Db	3302	CTTGAGGATGTTGTTCTGAGGCTGATATCTTTGTCAACCACCACCGTAAACAAGACAT	3361
QY	843	AGTGACACGGGAGCAGCTGGATCGCATGAAAAACAGTTGTATCGTATGCAATATGGCCA	902
Db	3362	CATCATGGTTGACCCACATGAGGAATAATGAAGAACAATGCCATGTTTGCAACATTGGTCA	3421
QY	903	CTCCACACAGAAATCGATGTGACCAAGCCTCCGCACT---CCGGAGCTGACGTGGGAGCG	959
Db	3422	CTTTGACAATGAGATCGACATGCTTGGGCTGGAGAACTACCCCGGCGTGAAGCGCATCAC	3481

QY	960	AGTACGTTCTCAGGTGGACCATGTCATCTGGCCAGATGGCCAAACGAGTTGTC---CTCCT	1016
Db	3482	CATCAAGCCCCCAAACTGACAGATGGGTCTTCCCTGGAGACCAACACCGGTATCATGTGCTT	3541
QY	1017	GCGCAGAGGGTCGTCTACTCAATTGGAGCTGCTCCACAGTTCCACC--TTTGTTCTGTGTC	1073
Db	3542	GGCTGAGGGTCGATGTGAACTTGGGATGCGCCACTGGACACCCCAAGTTTGTGATGTC	3601
QY	1074	CATCACAGCCACAACACAGGCTTTGGCAGTGAAGTAAGTCTATAATGCACCCG---AGGG	1130
Db	3602	CTGCTCCTTACCACACAGGTCATGTGCTCAGCTTGAAGTGTGGAAGAGAGAGTACCGG	3661
QY	1131	GCGATACAAGCAGGATGTGTACTTGCCTTCCTAAGAAAATGAGTGAATACGTTGCCAGCTT	1190
Db	3662	CAAGTACGAGAAGAAGGTTTACGTTTGGCCCAAGCACCCTTGATGAGAAGGTGGCTGCACCT	3721
QY	1191	GCACTCTGCCATCATTTGATGCCCACTTACAGAGCTGACAGATGACCAAGCAAAATATCT	1250
Db	3722	TCACCTGGGCAAACTTGGAGCTTAAGCTGACCCAGCCTTAAGCAAGTCCCAAGGCTGATTAAT	3781
QY	1251	GCGACCTCAACAAAAATGGGCCATTCAAACCTAATTATTACAGATAC	1296
Db	3782	CAGTGTGCCTGTTGAGGGTCCATACAAAGCCTGCTCACTACAGGTAC	3827

RESULT 14

```

US-09-754-853A-4/c
; Sequence 4, Application US/09754853A
; Publication No. US20030005491A1
;
; GENERAL INFORMATION:
; APPLICANT: Hauge, Brian M.
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Parsons, Jeremy D.
; APPLICANT: Wang, Ming Li
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: Soybean Cyst Nematode Resistance
; FILE REFERENCE: 38-10(15810)B
; CURRENT APPLICATION NUMBER: US/09/754, 853A
; CURRENT FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 60/174, 880
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 1119
;
; SEQ ID NO 4
; LENGTH: 513509
;
; TYPE: DNA
;
; ORGANISM: Glycine max
;
; FEATURE:
;
; NAME/KEY: CDS
; LOCATION: (111805)..(113968),(114684)..(115204)
;
; FEATURE:
;
; NAME/KEY: unsure
; LOCATION: (1)..(513509)
;
; OTHER INFORMATION: unsure at all n locations
;
; OTHER INFORMATION: Clone ID: 318013_region_A3
;
US-09-754-853A-4

```

Query Match	13.3%;	Score 172.4;	DB 9;	Length 513509;
Best Local Similarity	55.28;	Pred. No. 2.5e-46;		
Matches 423; Conservative	0;	Mismatches 331;	Indels 12;	Gaps 4;

[illegible]

```
Db 178798 GACCGAGATTGATCCCATCTGTGCCCTCAGGCTCTCATGGAAGGCTTCAGGTCTGTAC 178739
QY 783 GCTAAATGAGTCAATCCGGCAGTGCATGCTGTAATTAATCTGCACAGCAATAAGATGT 842
Db 178738 CTGGAGAGATGTTGTTCTGTAGGCTGATATCTTTGTCACCACCACCGGTAAACAAGACAT 178679
QY 843 AGTGACACGGGAGGACTTGTGATCGCATGAAAAACAGTTGTATCGTATGCAATATGGCCA 902
Db 178678 CATCATGTTGACACCATGAGGAAATGAGAAACAATGCCATGTTTGCAACATTTGTCA 178619
QY 903 CTCCAACACAGAAATCGATGTGACCAAGCTCCGCACT--CCGAGCTGACGTGGAGCG 959
Db 178618 CTTTGACAAATGAGATCGACATGCTTGGGCTGAGAACTACCCCGGCTGAAAGCGCATCAC 178559
QY 960 AGTACGTTCTCAGGTGAGACCATGTCACTGGCCAGATGGCAACGAGTTGTC--CTCCT 1016
Db 178558 CATCAAGCCCCAAACTGACAGATGGGTCTTCCTGAGACCAACACCGGTATCATTTGCTT 178499
QY 1017 GGCAGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTTCACCACC--TTGTCTGTGTC 1073
Db 178498 GGCAGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTTCACCACC--TTGTCTGTGTC 178439
QY 1074 CATCACAGCCACACACAGGCTTTGGCAGCTGATGAACTCTATATGCAACCCG--AGGG 1130
Db 178438 CTGCTCCTTCAACCAACCAAGGTCATGCTCAGCTTGAAGTGTGGAAGGAGAGATACCGG 178379
QY 1131 GCGATACAAGCAGGATGTGTACTTGTCTTCTAAGAAAAATGATGATACGTTGCCAGCTT 1190
Db 178378 CAAGTACGAGAGAAAGGTTTACGTTTGGCCAGCAACCTTGATGAGAAAGGTGGCTGCACT 178319
QY 1191 GCATCTGCCATCAATTGATGCCCACTTACAGAGCTGACAGATGACCAAGCAAAATATCT 1250
Db 178318 TCACCTGGGCAAACTTGAGCTAAGCTGACCACCTTAGCAAGTCCAGGCTGATTACAT 178259
QY 1251 GGGACTCAACAAAAATGGGCACTTCAAACTAATATTACAGATAC 1296
Db 178258 CAGTGTCCCTGTGAGGGTCCATACAGCCCTGCTCACTACAGGTAC 178213
```

RESULT 15

```
US-09-960-352-4421
; Sequence 4421, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathalagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 4421
; LENGTH: 433
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 19-LIB34-025-Q1-E1-E3
US-09-960-352-4421
```

```
Query Match 13.0%; Score 168.2; DB 10; Length 433;
Best Local Similarity 64.0%; Pred. No. 8.9e-47;
Matches 254; Conservative 0; Mismatches 143; Indels 0; Gaps 0;
```

```
QY 193 GAGACACTCTGTGCTGGGGCTCAGTCCGCTGCTGCTGTGTAACATCTACTCAACT 252
Db 37 GAGACCTCTGTTGCTGGGCTGAGGTGCGGTGCTCCAGCTGCAATATCTTCTCCACC 96
QY 253 CAGATGAGTAGCTGAGCAGCACTGGCTGAGGCTGAGCTTGCAGTGTTCGCTTGGAAAGGC 312
Db 97 CAGGACCATGACAGCTGCTCCATTGCAAGGCTGAGCTTCCAGTGTACGCGCTGGAAAGGT 156
```

```
QY 313 GAGTCAGAGATGACTTCTGTGCTGTATTCACCGCTGTGTGAACATGATGGGTGCAG 372
Db 157 GAACCGGATGAGAGTACCTGTGGTGCATTCAGACAGACGCTGTACTTCAAGGAGCGGCC 216
QY 373 GCCAATGATCTCTGATGATGGGGGAGACTTAACCCACTGGGTTTATAGAGTATCCA 432
Db 217 CTCACATGATCTTGAGCAGCAGCGGTGTGACTCACCACCTCATCCACACAGTACCCG 276
QY 433 AACGTGTTAAGAGATCCGAGGCAATGTGAGAGAGAGCGTGAAGTGTGTCACAGGCTG 492
Db 277 CAGCTCTGTGAGGCAATCCGAGGCAATCTGAGAGAGACCAACAGGGGGTCCACAACTG 336
QY 493 TATCAGCTCTCCAAAGCTGGGAAGCTCTGTGTTCCGGCCATGAACGTCAATGATCTGTT 552
Db 337 TACAAGATGATGGCCAAAGGATCCTGAAGGTGCCGGCCATCAATGTCAATGACTCTGTC 396
QY 553 ACCAAGAGAGATTGATACTTGTACTGTGCTGCCGAG 589
Db 397 ACCAAGAGAGATTGATACTTGTACTGTGCTGCCGAG 433
```

Search completed: April 21, 2003, 07:31:10
Job time : 854.419 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2003, 23:31:25 ; Search time 1749.62 Seconds
(without alignments)
11996.532 Million cell updates/sec

Title: US-09-782-051-1_COPY_549_1844

Perfect score: 1296
Sequence: 1 tccaagggcagcagcaattt.....aacctaattattacagatac 1296

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:*
- 1: em_estbda:*
 - 2: em_esthum:*
 - 3: em_estln:*
 - 4: em_estmu:*
 - 5: em_estov:*
 - 6: em_estpl:*
 - 7: em_estro:*
 - 8: em_htc:*
 - 9: gb_est1:*
 - 10: gb_est2:*
 - 11: gb_htc:*
 - 12: gb_est3:*
 - 13: gb_est4:*
 - 14: gb_est5:*
 - 15: em_estfun:*
 - 16: em_estom:*
 - 17: gb_gss:*
 - 18: em_gss_hum:*
 - 19: em_gss_inv:*
 - 20: em_gss_pln:*
 - 21: em_gss_vrt:*
 - 22: em_gss_fun:*
 - 23: em_gss_mam:*
 - 24: em_gss_mus:*
 - 25: em_gss_other:*
 - 26: em_gss_pro:*
 - 27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	845.6	65.2	1042	13	BM449470	BM449470 AGENCOURT
2	834.8	64.4	3234	11	AK014539	AK014539 Mus muscu
3	775.6	59.8	871	12	BG708016	BG708016 602671341
4	732.2	56.5	894	14	BQ896777	BQ896777 AGENCOURT
5	730.4	56.4	748	9	AU124992	AU124992 AU124992
6	720.4	55.6	933	14	BQ889173	BQ889173 AGENCOURT

7	716.4	55.3	747	12	BG697210	BG697210 602660444
8	716.4	55.3	832	10	BE270773	BE270773 600943931
9	698.2	53.9	751	9	AU133711	AU133711 AU133711
10	692.4	53.4	745	9	AU131962	AU131962 AU131962
11	687.8	53.1	691	14	BM837838	BM837838 K-EST0113
12	673	51.9	1098	13	BM454352	BM454352 AGENCOURT
13	672.4	51.9	956	12	BF792750	BF792750 602253481
14	666.2	51.4	938	9	AL527928	AL527928 AL527928
15	661.8	51.1	930	12	BG751929	BG751929 602731492
16	655.8	50.6	764	9	AU126038	AU126038 AU126038
17	648.4	50.0	673	12	BG501360	BG501360 602547990
18	644	49.7	1379	14	U69200	U69200 U69200 Soar
19	635.8	49.1	773	14	BQ745712	BQ745712 UT-M-EMO-
20	629.8	48.6	927	10	BE560874	BE560874 601346168
21	627.4	48.4	629	14	BM786844	BM786844 K-EST0065
22	622.6	48.0	827	12	BE777877	BE777877 601463641
23	621	47.9	839	14	BQ212518	BQ212518 AGENCOURT
24	613	47.3	614	14	BM837761	BM837761 K-EST0113
25	603.8	46.6	819	10	BE542285	BE542285 601067027
26	599.8	46.3	969	12	BG259448	BG259448 602378544
27	597.6	46.1	639	10	AW958506	AW958506 EST370576
28	593.8	45.8	643	14	BM995818	BM995818 UT-H-ED0-
29	592	45.7	598	14	BO614546	BO614546 1K9d02.Y
30	588.8	45.4	664	10	AV751644	AV751644 AV751644
31	588.8	45.4	947	12	BF792477	BF792477 602353566
32	587.6	45.3	1124	14	BQ062203	BQ062203 AGENCOURT
33	586.4	45.2	867	13	BM453920	BM453920 AGENCOURT
34	585	45.1	1021	14	BQ059373	BQ059373 AGENCOURT
35	582.2	44.9	1002	12	BE900605	BE900605 601673502
36	578	44.6	590	10	BE514092	BE514092 601315914
37	575.6	44.4	665	12	BG898694	BG898694 HOA4-1-A-
38	572.6	44.2	672	12	BF342864	BF342864 602015152
39	565	43.6	844	10	BE336849	BE336849 601064852
40	564.4	43.5	731	12	BG826655	BG826655 602748977
41	562	43.4	700	9	AU133884	AU133884 AU133884
42	561.4	43.3	975	10	BE561484	BE561484 601346553
43	558.8	43.1	603	12	BG244104	BG244104 602358170
44	551.6	42.6	616	12	BF783516	BF783516 602111292
45	542.2	41.8	858	9	AL554850	AL554850 AL554850

ALIGNMENTS

RESULT 1
BM449470
LOCUS
DEFINITION
AGENCOURT_6400881 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5493677
5', mRNA sequence.
ACCESSION
BM449470
VERSION
BM449470.1 GI:18498510
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
NIH-MGC <http://mgc.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LLM12117 row: m column: 06
High quality sequence stop: 694.
Location/Qualifiers
1. 1042

FEATURES
source

OY	1198	CCAATCATTTGATGCCACCTTACAGAGCTGCACAGATGACCACAAGCAATAATCTGGGACTC	1257
Dd	1787	CCCCACCTTTGATGCCACCTGACAGAGCTAACAGATGAACAAGCCAAAGTATCTGGGACTC	1846
OY	1258	AACAAAAATGGGGCATTCAAACCTTAATTATTACAGATA	1295
Dd	1847	AACAAGAATGGACCGCTTCAAGCCTTAAGCTAAGGTGTA	1884
RESULT 3			
LOCUS	BG708016	871 bp	mRNA linear EST 07-MAY-2001
DEFINITION	602671341F1 NIH_MGC_96 Homo sapiens CDNA clone IMAGE:4793770 5',		
ACCESSION	BG708016		
VERSION	BG708016.1	GI:13984941	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 871)		
JOURNAL	NIH-MGC http://mgs.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC) unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabps-remail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN) CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov plate: LLAM10674 row: b column: 11 High quality sequence stop: 839.		
FEATURES	location/Qualifiers		
source	1..871		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4793770"		
	/clone_lib="NIH_MGC_96"		
	/tissue_type="hypothalamus"		
	/lab_host="DH10B"		
	/note="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTVN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, In preparation). Library constructed by M. Brownstein (NHGRI, National Institutes of Health). Note: this is a NIH_MGC library."		
BASE COUNT	238 a 190 c 235 g 207 t 1 others		
ORIGIN			
Query Match	59.8%; Score 775.6; DB 12; Length 871;		
Best Local Similarity	98.5%; Pred. NO. 1.9e-226;		
Matches 803; Conservative	0; Mismatches 10; Indels 2; Gaps 2;		
OY	360	GGATGGGTGGCAGGCGCACATGATCCTTGATGATGGGGGAGACTTAACCCACTGGGTTTA	419
Dd	5	GATGGGTGGCAGGCGCACATGATCCTTGATGATGGGGGAGACTTAACCCACTGGGTTTA	64
OY	420	TAGAAGTATCCAAACGCTGTTTAAGAAGATCCGAGGCATTGTGGAAGAGAGCGTGACTGG	479
Dd	65	TAGAAGTATCCAAACGCTGTTTAAGAAGATCCGAGGCATTGTGGAAGAGAGCGTGACTGG	124
OY	480	TGTTACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTGTTCCGGGCATGAACGT	539
Dd	125	TGTTACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTGTTCCGGGCATGAACGT	184

QY	540	CAATGATTTCTGTTACCAACAAGAGTTTGATTAACCTTGTACTGCTGCGGAGAAATCCATT	599
Db	185	CAATGATTTCTGTTACCAACAAGAGTTTGATTAACCTTGTACTGCTGCGGAGAAATCCATT	244
QY	600	GGATGGCCTGAAGAGGACCACAGATGTGATGTTTGGTGGGAACAAGTGGTGTGTGG	659
Db	245	GGATGGCCTGAAGAGGACCACAGATGTGATGTTTGGTGGGAACAAGTGGTGTGTGG	304
QY	660	CTATGCTGAGGTAGGCAAGGCGCTGCTGTGCTGCTCTCAAGCTCTTGAGCAATTGTCTA	719
Db	305	CTATGCTGAGGTAGGCAAGGCGCTGCTGTGCTGCTCTCAAGCTCTTGAGCAATTGTCTA	364
QY	720	CATTACCGAAATCGACCCCATCTGTGCTCTGCGAGGCGCTGCATGATGGGTTCAGGGTGT	779
Db	365	CATTACCGAAATCGACCCCATCTGTGCTCTGCGAGGCGCTGCATGATGGGTTCAGGGTGT	424
QY	780	AAAGCTAATGAAGTCATCCGGCAAGTCGATGTGCTAATAACTTGACACAGGAATAAGAA	839
Db	425	AAAGCTAATGAAGTCATCCGGCAAGTCGATGTGCTAATAACTTGACACAGGAATAAGAA	484
QY	840	TGTAGTGCACACGGGAGCACCTTGGATCCGATGAAAAACAGTTGTATCGTATGCAATATGG	899
Db	485	TGTAGTGCACACGGGAGCACCTTGGATCCGATGAAAAACAGTTGTATCGTATGCAATATGG	544
QY	900	CCACTCCCAACACAGAATCGATGTGACCAAGCCTCCGCACCTCCGAGCTGACGTGGAGCG	959
Db	545	CCACTCCCAACACAGAATCGATGTGACCAAGCCTCCGCACCTCCGAGCTGACGTGGAGCG	604
QY	960	AGTACGTTCTCAGGTGACCCATGTTCATCTGGCCAGATGGCAACGAGTGTCTCTGCG	1019
Db	605	AGTACGTTCTCAGGTGACCCATGTTCATCTGGCCAGATGGCAACGAGTGTCTCTGCG	664
QY	1020	AGAGGGTCTGCTACTCAATTTGAGCTGCTCCACAGTTCCACACTTTGTCTGTCCATCAC	1079
Db	665	AGAGGGTCTGCTACTCAATTTGAGCTGCTCCACAGTTCCACACTTTGTCTGTCCATCAC	724
QY	1080	AGCCACAACACAGGCTTTGGCACT-GATAGAACTCTATAATGCAACCCGAGGGCGGATACA	1138
Db	725	AGNCACAACACAGGCTTTGGCACTGATAGAACTCTATAATGCAACCCGA-GGGCGATACA	783
QY	1139	AGCAGGATGTGTACTTGTCTCTTAAGAAAAATGGAT 1173	
Db	784	AGCAGGATGTGTACTTGTCTCTTAAGAAAAATGGAT 818	

RESULT 4
BQ896777
LOCUS
DEFINITION
BQ896777 894 bp mRNA linear EST 16-AUG-2002
AGENCOURT_8715624 NIH_MGC_101 Homo sapiens CDNA clone IMAGE:6425821
5', mRNA sequence.
ACCESSION
BQ896777
VERSION
BQ896777.1 GI:22288791
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 894)
NIH-MGC <http://mgc.ncl.nih.gov/>.
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE
Unpublished (1999)
JOURNAL
Contact: Robert Strausberg, Ph.D.
COMMENT
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
plate: LICM2608 row: d column: 14
high quality sequence stop: 700.

FEATURES
source

Location/Qualifiers
1. 894
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6425821"
/clone_1lb="NIH_MGC_101"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site_1: EcoRI; Site_2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

BASE COUNT 226 a 205 c 247 g 216 t
ORIGIN

Query Match 56.5%; Score 732.2; DB 14; Length 894;
Best Local Similarity 99.3%; Pred. No. 4.2e-213;
Matches 746; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

OY 362 ATGGGTGGCAGGCCAATCATGATCCTGGATGATGGGGAGAGACTTAACCCACTGGGTTTATA 421
DB 1 ATGGGTGGCAGGCCAATCATGATCCTGGATGATGGGGAGAGACTTAACCCACTGGGTTTATA 60
OY 422 AGAGTATCCAAAGCTGTTTAAAGATCCGAGCATTTGTGAAGAGAGAGCGTGTGCTG 481
DB 61 AGAGTATCCAAAGCTGTTTAAAGATCCGAGCATTTGTGAAGAGAGAGCGTGTGCTG 120
OY 482 TTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTCTCCGGCATGAGCTCA 541
DB 121 TTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTCTCCGGCATGAGCTCA 180
OY 542 ATGATTTCTGTACCAACAGAAAGTTGATTAATCTGTACTGTCTGCCGAGAAATTCATTTTG 601
DB 181 ATGATTTCTGTACCAACAGAAAGTTGATTAATCTGTACTGTCTGCCGAGAAATTCATTTTG 240
OY 602 ATGGCTGAAGAGACACAGATGTGTGTTGGTGGGAAACAAGTGGTGTGTGGCT 661
DB 241 ATGGCTGAAGAGACACAGATGTGTGTTGGTGGGAAACAAGTGGTGTGTGGCT 300
OY 662 ATGGTGAAGTGAAGAGGCTGTGTGTCTCTCAAGCTCTTGAGCAATGTCTTACA 721
DB 301 ATGGTGAAGTGAAGAGGCTGTGTGTCTCTCAAGCTCTTGAGCAATGTCTTACA 360
OY 722 TTACCGAAATCGACCCATCTGTCTGTGACAGGCTGCATGATGGGTTGAGGTTGATA 781
DB 361 TTACCGAAATCGACCCATCTGTCTGTGACAGGCTGCATGATGGGTTGAGGTTGATA 420
OY 782 AGCTAATGAAGTATCCGGCAAGTGCATGTCTGTAATACTTGACACAGAAATAGAATG 841
DB 421 AGCTAATGAAGTATCCGGCAAGTGCATGTCTGTAATACTTGACACAGAAATAGAATG 480
OY 842 TAGTGACACGGGAGCACTTGATCGCATGAAAAACAGTTGTATCGTATGCAATATGGCC 901
DB 481 TAGTGACACGGGAGCACTTGATCGCATGAAAAACAGTTGTATCGTATGCAATATGGCC 540
OY 902 ACTCCAAACAGAAATCGATGTGACCAAGCTCCGCACTCCGAGCTGACGTGGAGCGAG 961
DB 541 ACTCCAAACAGAAATCGATGTGACCAAGCTCCGCACTCCGAGCTGACGTGGAGCGAG 600
OY 962 TAGTTCTCAGGTGAGCATGTCTGAGCAGATGGCAAAAGAGTTGCTCTCTGTCAG 1021
DB 601 TAGTTCTCAGGTGAGCATGTCTGAGCAGATGGCAAAAGAGTTGCTCTCTGTCAG 660
OY 1022 AGGGTCTCTACTCAATTTGAGCTGCTCACAGTTCCACCTTTGTTCTGTCAATCAG 1081
DB 661 AGGGTCTCTACTCAATTTGAGCTGCTCACAGTTCCACCTTTGTTCTGTCAATCAG 720
OY 1082 CCACACACAGG--CTTGGCACTGATAGAA 1110
|||||

DB 721 CCACACACAGGCTTTTGCCACTGATAGAA 751

RESULT 5
AU124992
LOCUS
DEFINITION AU124992 NT2RM4 Homo sapiens cDNA clone NT2RM4000853 5', mRNA
sequence.
ACCESSION AU124992
VERSION AU124992
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

748 bp
mRNA
linear
EST
01-AUG-2002
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 748)
Ota, T., Wakamatsu, A., Ozawa, M., Ishii, S., Saito, K., Yamamoto, J.,
Nakamura, Y., Nishikawa, T., Nagai, T., Suzuki, Y., Sugano, S. and
Isogai, T.
HRI human cDNA project (Ota, T., Wakamatsu, A., Ozawa, M., Ishii, S.,
Saito, K., Yamamoto, J., Nakamura, Y., Nishikawa, T., Nagai, T., Suzuki,
Y., Sugano, S., Isogai, T.)
Unpublished (2000)
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.

FEATURES
source
Location/Qualifiers
1. 748
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RM4000853"
/clone_1lb="NT2RM4"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/note="Vector: pME18SFL3; mRNA from uninduced NT2 neuronal
precursor cells"

BASE COUNT 181 a 152 c 226 g 185 t 4 others
ORIGIN
Query Match 56.4%; Score 730.4; DB 9; Length 748;
Best Local Similarity 99.2%; Pred. No. 1.3e-212;
Matches 742; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

OY 61 GAGATTGAGATTGACAGAGACATGTCTGCTGTGATTTCACTCAGAAACGTCCTCAG 120
DB 1 GAGATTGAGATTGACAGAGACATGTCTGCTGTGATTTCACTCAGAAACGTCCTCAG 60
OY 121 GGGGAGAAGCCCTGGCTGCTTAAATAGTGGGCTGTACACACATCAGCCAGACA 180
DB 61 GGGGAGAAGCCCTGGCTGCTTAAATAGTGGGCTGTACACACATCAGCCAGACA 120
OY 181 GCGGTGTGATTGAGACACTGTGCCCCGCGGCTCAGTCCGCTGCTGCTGTAAC 240
DB 121 GCGGTGTGATTGAGACACTGTGCCCCGCGGCTCAGTCCGCTGCTGCTGTAAC 180
OY 241 ATCTACTCAACTCAGAAATGAGTACGTGACGACACTGGCTGAGGCTGAGTTGACGTGTC 300
DB 181 ATCTACTCAACTCAGAAATGAGTACGTGACGACACTGGCTGAGGCTGAGTTGACGTGTC 240
OY 301 GCTTGAAGGGGAGTCAAGATGACTTCTGTGTGTATTTGACCGCTGTGTGACATG 360
DB 241 GCTTGAAGGGGAGTCAAGATGACTTCTGTGTGTATTTGACCGCTGTGTGACATG 300
OY 361 GATGGGTGGCAGGCCAATGATCCTGTGATGGGAGAGACTTAACCCACTGGGTTAT 420
|||||

Db 301 GATGGTGGCAGGCCACATGATCTGTGATGAGGGGAGACTTAACCCACTGGTTTAT 360
OY 421 AAGAAGTATCCAAAGCTGTTTAAAGATCCGAGCATTTGTGAAGAGAGAGCGTGAAGT 480
Db 361 AAGAAGTATCCAAAGCTGTTTAAAGATCCGAGCATTTGTGAAGAGAGAGCGTGAAGT 420
OY 481 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGAGACTCTGTGTTCCGGCAATGAACGTC 540
Db 421 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGAGACTCTGTGTTCCGGCAATGAACGTC 480
OY 541 AATGATCTCTTACCAACAGAGATTGATTAATCTTACTGCTGCTGCCAGAAATTCATTG 600
Db 481 AATGATCTCTTACCAACAGAGATTGATTAATCTTACTGCTGCTGCCAGAAATTCATTG 540
OY 601 GATGGCTGAAGAGACACAGATGTGATGTTGGTGGGAAACAAGTGGTGGTGGTGGC 660
Db 541 GATGGCTGAAGAGACACAGATGTGATGTTGGTGGGAAACAAGTGGTGGTGGTGGC 600
OY 661 TATGTTAGGTAGGCAAGGGCTGCTGCTGCTCTCAAGCTCTTGAGCAATGCTCTAC 720
Db 601 TATGTTAGGTAGGCAAGGGCTGCTGCTGCTCTCAAGCTCTTGAGCAATGCTCTAC 660
OY 721 ATTA-CCGAATGCAAGCTGCTGCTGCTGCTGCAAGCTGCTGATGATGGTTCAGGCTGT 779
Db 661 ATTACCCGAATGCAAGCTGCTGCTGCTGCTGCAAGCTGCTGATGATGGTTCAGGCTGT 720
OY 780 AAGCTAATGAATGATCCTCCGCAAGTC 807
Db 721 AANCTAATGAATGATCCTCCGCAAGTC 748

RESULT 6
LOCUS BQ889173 933 bp mRNA linear EST 16-AUG-2002
DEFINITION AGENCOURT_8482604 lupskl_dorsal_root_ganglion Homo sapiens CDNA
clone IMAGE:6185747 5', mRNA sequence.
ACCESSION BQ889173
VERSION BQ889173.1 GI:22281187
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 933)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM13577 row: a column: 12
High quality sequence stop: 656.

FEATURES
source location/Qualifiers
1. 933
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6185747"
/clone_lib="lupskl_dorsal_root_ganglion"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6 (Life Technologies); Site_1:
NotI, Site_2: SalI; CDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGCGTCCG-3' and
5'-GACTAGTCTAGATCGCGAGCGGCCCT(15)-3'. Size selected >

1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies.
BASE COUNT 231 a 202 c 281 g 218 t 1 others
ORIGIN

Query Match 55.6%; Score 720.4; DB 14; Length 933;
Best Local Similarity 97.9%; Pred. No. 1.8e-209;
Matches 762; Conservative 0; Mismatches 11; Indels 5; Gaps 3;

OY 113 GTGCTCAGGGGAGAGAGCCCTGGCTGCTGCTAAATAGTGGGCTGTACACACATCACAG 172
Db 1 GTGCTCAGGGGAGAGAGCCCTGGCTGCTGCTAAATAGTGGGCTGTACACACATCACAG 60
OY 173 CCCAGACAGCGGTGTTGATGAGACACTCTGTGCTGGGGCTCAGTGGCGTGGTCTG 232
Db 61 CCCAGACAGCGGTGTTGATGAGACACTCTGTGCTGGGGCTCAGTGGCGTGGTCTG 120
OY 233 CTTGTACATCTACTCAACTCAGATGAGTAGTACGACACTGGCTGAGGCTGAGTGTG 292
Db 121 CTTGTACATCTACTCAACTCAGATGAGTAGTACGACACTGGCTGAGGCTGAGTGTG 180
OY 293 CAGTGTGCTTGAAGGGCGAGTCAGAGATGACTTCTGTGTGTATGACCGCTGTG 352
Db 181 CAGTGTGCTTGAAGGGCGAGTCAGAGATGACTTCTGTGTGTATGACCGCTGTG 240
OY 353 TGAACATGATGGGTGGCAGGCCCAACATGATCCTGATGATGGGGAGACTTAACCACT 412
Db 241 TGAACATGATGGGTGGCAGGCCCAACATGATCCTGATGATGGGGAGACTTAACCACT 300
OY 413 GGGTTTATAGAAGTATCCAAACGCTGTTAAGAGATCCGAGGCTTGTGGAAGAGAGCG 472
Db 301 GGGTTTATAGAAGTATCCAAACGCTGTTAAGAGATCCGAGGCTTGTGGAAGAGAGCG 360
OY 473 TGACTGGTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTGTTCCGCCA 532
Db 361 TGACTGGTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTGTTCCGCCA 420
OY 533 TGAACGTCAATGATCTGTTACCAACAGAGATTGATACTGTACTGCTGCCGAGAAT 592
Db 421 TGAACGTCAATGATCTGTTACCAACAGAGATTGATACTGTACTGCTGCCGAGAAT 480
OY 593 CCATTTGGATGGCTGAAGAGGACACAGATGATGTTGGTGGGAACAAGTGTGG 652
Db 481 CCATTTGGATGGCTGAAGAGGACACAGATGATGTTGGTGGGAACAAGTGTGG 540
OY 653 TGTGTGCTATGTTAGGTAGGCAAGGGCTGCTGCTCTCAAGCTCTTGGAGCAA 712
Db 541 TGTGTGCTATGTTAGGTAGGCAAGGGCTGCTGCTCTCAAGCTCTTGGAGCAA 600
OY 713 TTGCTACATTAACGAATCGACCCCATCTGCTCTGACGGCTGCATGATGGG--TT 770
Db 601 TTGCTACATTAACGAATCGACCCCATCTGCTCTGACGGCTGCATGATGGGGTTC 660
OY 771 CAGGTTGTAAGCTAAATGAAGTATCCGCAAGTGCATGCTGTAATTAATTGACAGG 830
Db 661 AGGGTGGTAAAGCTAAATGAAGTATCCGCAAGTGCATGCTGTAATTAATTGACAGG 720
OY 831 AATAAGAAATGATGACAC--GGGAGCACTTGGATCGCATG--AAAAAGTTGTATC 885
Db 721 AATAAGAAATGATGACACCGGGAGCACTTGGATCGCATGCAAAAACAGTTGTATC 778

RESULT 7
LOCUS BG697210 747 bp mRNA linear EST 07-MAY-2001
DEFINITION 60266044F1 NCI-CGAP_Skn3 Homo sapiens CDNA clone IMAGE:4803515 5',
mRNA sequence.
ACCESSION BG697210
VERSION BG697210.1 GI:13963166
KEYWORDS EST.

SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 747)
 NIH-MGC http://mgi.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: James Cleaver, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Plate: L1AM10699 row: h column: 12
 High quality sequence stop: 747.
 Location/Qualifiers

FEATURES
 source 1..747
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4803515"
 /clone_1lb="NCI_CGAP_Skn3"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: Skin; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.5kb. Library constructed by Life Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 180 a 152 c 232 g 183 t
 ORIGIN

Query Match 55.3%; Score 716.4; DB 12; Length 747;
 Best Local Similarity 99.6%; Pred. No. 2.6e-208;
 Matches 739; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 34 ATCAAGCAGCAGAAATTGGACGCCGGAGATTGAGATTGCAGAGCAAGACATGTCCT 93
 Db 1 ATCAAGCAGCAGAA-TTGGACGCCGGAGATTGAGATTGCAGAGCAAGACATGTCCT 59
 QY 94 CTGATTTCACTCAGGAACGTCCTCAGGGGAGAAAGCCCTGGCTGGCTAAATAATAGTG 153
 Db 60 CTGATTTCACTCAGGAACGTCCTCAGGGGAGAAAGCCCTGGCTGGCTAAATAATAGTG 119
 QY 154 GCGTGTACACATCACAGCCAGACAGCGGTGTGATTGAGACACTGTGCTGGCTGGGG 213
 Db 120 GCGTGTACACATCACAGCCAGACAGCGGTGTGATTGAGACACTGTGCTGGGG 179
 QY 214 GCTCAGTCCGCTGTGCTGTGAACATCTACTCAACTCAGATGAGATGAGCTGCAGCA 273
 Db 180 GCTCAGTCCGCTGTGCTGTGAACATCTACTCAACTCAGATGAGATGAGCTGCAGCA 239
 QY 274 CTGGCTGAGGCTGAGTTGACAGTGTTCGCTTGAAGGGCGAGTCAGAGATGACTTCTGG 333
 Db 240 CTGGCTGAGGCTGAGTTGACAGTGTTCGCTTGAAGGGCGAGTCAGAGATGACTTCTGG 299
 QY 334 TGGTGTATTGACCGCTGTGTGAACATGATGGGTGGCAGGCCAATGATCTCTGATGAT 393
 Db 300 TGGTGTATTGACCGCTGTGTGAACATGATGGGTGGCAGGCCAATGATCTCTGATGAT 359
 QY 394 GGGGAGACTTAACCCACTGGGTTTAAAGAGTATCCAAA-CGTGTTTAAAGAGATCCG 452
 Db 360 GGGGAGACTTAACCCACTGGGTTTAAAGAGTATCCAAAAGCGTTTAAAGAGATCCG 419
 QY 453 AGGCATGTGGAAGAGAGCGTGTGATGCTTCAAGGCTGTATCAGCTCTCCAAAGCTGG 512
 Db 420 AGGCATGTGGAAGAGAGCGTGTGATGCTTCAAGGCTGTATCAGCTCTCCAAAGCTGG 479
 QY 513 GAAGCTCTGTGTCCGGCATGAAGCTCAATGATTCTGTACCAAGAAGATTGTGATAA 572
 Db 480 GAAGCTCTGTGTCCGGCATGAAGCTCAATGATTCTGTACCAAGAAGATTGTGATAA 539

QY 573 CTGTACTGCTGCCAGAAATCCATTTTGATGGCCCTGAGAGGACCACAGATGTGATGT 632
 Db 540 CTGTACTGCTGCCAGAAATCCATTTTGATGGCCCTGAGAGGACCACAGATGTGATGT 599
 QY 633 TGGTGGAAACAAGTGTGTGTGTGCGCTATGATGATGAGTACGCAAGGCTGTGTGTC 692
 Db 600 TGGTGGAAACAAGTGTGTGTGTGCGCTATGATGATGAGTACGCAAGGCTGTGTGTC 659
 QY 693 TCTCAAGCTCTTGAGCAATTGTCTACATTACCGAATCGAACCCTATCTGTCTGTGCA 752
 Db 660 TCTCAAGCTCTTGAGCAATTGTCTACATTACCGAATCGAACCCTATCTGTCTGTGCA 719
 QY 753 GGCCTGCATGATGGGTTCAGG 774
 Db 720 GGCCTGCATGATGGGTTCAGG 741

RESULT 8
 BE270773 832 bp mRNA linear EST 13-JUL-2000
 LOCUS BE270773
 DEFINITION 600943931F1 NIH_MGC_8 Homo sapiens CDNA clone IMAGE:2966598 5', mRNA sequence.
 ACCESSION BE270773
 VERSION BE270773
 KEYWORDS BE270773.1 GI:9144416
 SOURCE EST.
 ORGANISM human.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 832)
 NIH-MGC http://mgi.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: L1CM69 row: n column: 07
 High quality sequence start: 2
 High quality sequence stop: 767.
 Location/Qualifiers

FEATURES
 source 1..832
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2966598"
 /clone_1lb="NIH_MGC_8"
 /tissue_type="Burkitt lymphoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 236 a 192 c 202 g 201 t
 ORIGIN

Query Match 55.3%; Score 716.4; DB 10; Length 832;
 Best Local Similarity 99.5%; Pred. No. 2.8e-208;
 Matches 750; Conservative 0; Mismatches 1; Indels 3; Gaps 3;

QY 545 ATTCTGTACCAACAAGATTGATACTGTACTGCTGCCGAGAATCCATTGTGATG 604
 Db 13 ATTCTGTACCAACAAGATTGATACTGTACTGCTGCCGAGAATCCA-TTTGGATG 71
 QY 605 GCCTGAAGAGGACACAGATGTGATGTTGGTGGGAAACAAGTGGTGTGTGCTATG 664

```
Db 72 GCGTGAAGAGGACACAGATGTGATTTGGTGGAAACAGTGGTGTGTGGCTATG 131
QY 665 GTGAGTAGGCAAGGGCTGCTGCTGCTCTCAAGCTCTGGAGCAATGTCTACATTA 724
Db 132 GTGAGTAGGCAAGGGCTGCTGCTGCTCTCAAGCTCTGGAGCAATGTCTACATTA 191
QY 725 CCGAAATCGACCCCATCTGTGCTCTGACAGGCTGCATGATGGGTTCAGGGTGTAAAGC 784
Db 192 CCGAAATCGACCCCATCTGTGCTCTGACAGGCTGCATGATGGGTTCAGGGTGTAAAGC 251
QY 785 TAAATGAAGTCATCCGGCAAGTCGATGCTGTAATACTTCACAGGAATAAGAAATGTAG 844
Db 252 TAAATGAAGTCATCCGGCAAGTCGATGCTGTAATACTTCACAGGAATAAGAAATGTAG 311
QY 845 TGACACGGGAGCAGCTTGGATCGCATGAAACAGTGTGATCGTATGCAATATGGCCACT 904
Db 312 TGACACGGGAGCAGCTTGGATCGCATGAAACAGTGTGATCGTATGCAATATGGCCACT 371
QY 905 CCAACACAGAAATCGATGTGACCAAGCTCCGCACTCCGAGCTGACGTGGAGCGAGTAC 964
Db 372 CCAACACAGAAATCGATGTGACCAAGCTCCGCACTCCGAGCTGACGTGGAGCGAGTAC 431
QY 965 GTTCTCAGGTGGAACCATGTCTATGCGCCAGATGGCAACGAGTGTCTCTCCGACAGAG 1024
Db 432 GTTCTCAGGTGGAACCATGTCTATGCGCCAGATGGCAACGAGTGTCTCTCCGACAGAG 491
QY 1025 GTGCTCTACTCAATTTGAGCTGCTCCACAGTCCCACTTGTGTCTGTCATCAGACCA 1084
Db 492 GTGCTCTACTCAATTTGAGCTGCTCCACAGTCCCACTTGTGTCTGTCATCAGACCA 551
QY 1085 CAACACAGCTTTGGCAGCTGATAGAACTCTATATGCAACCCGAGGGCGATACAGCAGG 1144
Db 552 CAACACAGCTTTGGCAGCTGATAGAACTCTATATGCAACCCGAGGGCGATACAGCAGG 611
QY 1145 ATGTGTACTTCTCTCTTAAGAAATGATGATACGTTGCCAGCTTGCATCTGCCATCAT 1204
Db 612 ATGTGTACTTCTCTCTTAAGAAATGATGATACGTTGCCAGCTTGCATCTGCCATCAT 671
QY 1205 TTGATGCCCCACCTTACAGAGCTGACAGATGACCAAGC-AAATATCTGGGACTCAACAA 1263
Db 672 TTGATGCCCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAACAA 731
QY 1264 AATGGGCCATTC-AAACCTAATTATTACAGATAC 1296
Db 732 AATGGGCCATTCAAAACCTAATTATTACAGATAC 765

RESULT 9
LOCUS AU133711 751 bp mRNA linear EST 01-AUG-2002
DEFINITION AU133711 OVARC1 Homo sapiens cDNA clone OVARC1000494 5', mRNA
sequence.
ACCESSION AU133711
VERSION AU133711.1 GI:10994250
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 751)
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,Y., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Ishigaki,T.
TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
```

```
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
Location/Qualifiers
1. 751
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="OVARC1000494"
/clone_11b="OVARC1"
/tissue_type="ovary, tumor tissue"
/note="Vector: pME18SFL3"
BASE COUNT 192 a 149 c 222 g 184 t 4 others
ORIGIN
Query Match 53.9%; Score 698.2; DB 9; Length 751;
Best Local Similarity 99.0%; Pred. No. 1e-202;
Matches 711; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 101 CACTCAGAAACGTGCTCAGGGGAGAAAGCCCTTGGCTGTGCTAAATAGTGGCTGTA 160
Db 1 CACTCAGAAACGTGCTCAGGGGAGAAAGCCCTTGGCTGTGCTAAATAGTGGCTGTA 60
QY 161 CACACATCAGAGCCAGACAGAGCGGTGATGAGACACTCTGTGCTGGGGCTCAGT 220
Db 61 CACACATCAGAGCCAGACAGAGCGGTGATGAGACACTCTGTGCTGGGGCTCAGT 120
QY 221 GCCGCTGCTGCTGTACATCTACTCAACTCAGATGAGTAGCTGACGACTGGGTG 280
Db 121 GCCGCTGCTGCTGTACATCTACTCAACTCAGATGAGTAGCTGACGACTGGGTG 180
QY 281 AGGCTGAGTTCAGAGTGTTCGCTTGGAGGGCGAGTCAGAGATGACTTCTGTGTGTA 340
Db 181 AGGCTGAGTTCAGAGTGTTCGCTTGGAGGGCGAGTCAGAGATGACTTCTGTGTGTA 240
QY 341 TTGACCGCTGTGTAACATGATGGTGGAGGCCCAACATGATCCTGATGATGGGGAG 400
Db 241 TTGACCGCTGTGTAACATGATGGTGGAGGCCCAACATGATCCTGATGATGGGGAG 300
QY 401 ACTTAACCACTGGGTTTATAAGAGTATCCAAAGCTGTTTAAGAGATCCAGGCTATG 460
Db 301 ACTTAACCACTGGGTTTATAAGAGTATCCAAAGCTGTTTAAGAGATCCAGGCTATG 360
QY 461 TGAAGAGAGCGGTGACTGCTGTTCACAGGCTGTATCAGCTCTCCAAGCTGGGAAGCTCT 520
Db 361 TGAAGAGAGCGGTGACTGCTGTTCACAGGCTGTATCAGCTCTCCAAGCTGGGAAGCTCT 420
QY 521 GTGTTCCGGCCATGACGTCATGATCTGTACCAACAGAAAGTTGATTAAGTGTACT 580
Db 421 GTGTTCCGGCCATGACGTCATGATCTGTACCAACAGAAAGTTGATTAAGTGTACT 480
QY 581 GCTGCCGAGAAATCCATTTTGGATGGCCTGAAGAGACACAGATGTGATGTTGGTGGGA 640
Db 481 GCTGCCGAGAAATCCATTTTGGATGGCCTGAAGAGACACAGATGTGATGTTGGTGGGA 540
QY 641 AACAAAGTGTGTGTGCTATGATGAGGTAGGCAAGGGCTGCTGTCTCTCAAG 700
Db 541 AACAAAGTGTGTGTGCTATGATGAGGTAGGCAAGGGCTGCTGTCTCTCAAG 600
QY 701 CTCTTGAGCAATGTCTACATTAACGGAATTCAGCCCATCTGTCTCTGACAGGCTGCA 760
Db 601 CTCTTGAGCAATGTCTACATTAACGGAATTCAGCCCATCTGTCTCTGACAGGCTGCA 660
QY 761 TGGATGGGTTCAGGCT-GGTAAAGCTAAATGAAGTCATCCGGCAAGTGCATGCTGTA 817
Db 661 TGGATGGGTTCAGGCTGAGGTAAAGCTAAATGAAGTCATCCGGCAANTCCATGCTNTAA 718

RESULT 10
LOCUS AU131962 745 bp mRNA linear EST 01-AUG-2002
DEFINITION AU131962 NT2RP3 Homo sapiens cDNA clone NT2RP3003545 5', mRNA
sequence.
ACCESSION AU131962
```

VERSION AU131962.1 GI:10992316
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 745)
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,T.
TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix
Research Institute; cDNA library construction; Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
source
1. 745
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RP3003545"
/clone_1lb="NT2RP3"
/cell_type="teratocarcinoma"
/note="Vector: PME18SFL3; mRNA from NT2 neuronal precursor
cells after 2-weeks retinoic acid (RA) induction"
BASE COUNT 193 a 149 c 214 g 186 t 3 others
ORIGIN
Query Match 53.4%; Score 692.4; DB 9; Length 745;
Best Local Similarity 99.6%; Pred. No. 6e-201;
Matches 704; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 229 TCTGCTGTACATCTACTCACTCAGATGAGTGTGACGACCTGGCTGAGGCTGGA 288
|||||
DB 1 TCTGCTGTACATCTACTCACTCAGATGAGTGTGACGACCTGGCTGAGGCTGGA 60
QY 289 GTTGACAGTGTTCGTTGGAAGGGCGAGTCAGAGATGACTTCTGTTGTTATGACCGC 348
|||||
DB 61 GTTGACAGTGTTCGTTGGAAGGGCGAGTCAGAGATGACTTCTGTTGTTATGACCGC 120
QY 349 TGTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 408
|||||
DB 121 TGTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
QY 409 CACTGGGTTTATAAGATATCCAAACGTTTAAAGATCCGAGGATGTGGAAGAG 468
|||||
DB 181 CACTGGGTTTATAAGATATCCAAACGTTTAAAGATCCGAGGATGTGGAAGAG 240
QY 469 AGCGTACTGTTGTTACAGAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTTCCG 528
|||||
DB 241 AGCGTACTGTTGTTACAGAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTTCCG 300
QY 529 GCCATGACGTCATGATCTCTGTACCAACAGAGTTTGTATTAATCTGTACTGCTGCCGA 588
|||||
DB 301 GCCATGACGTCATGATCTCTGTACCAACAGAGTTTGTATTAATCTGTACTGCTGCCGA 360
QY 589 GAATCCATTTTGGATGGCTGAAGAGGACACAGATGTGATGTTTGGTGGGAACAAGTG 648
|||||
DB 361 GAATCCATTTTGGATGGCTGAAGAGGACACAGATGTGATGTTTGGTGGGAACAAGTG 420
QY 649 GTGTGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 708
|||||
DB 421 GTGTGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 480
QY 709 GCAATGTCTACATTAACCAATGACCCCAATCTGTCTCTGACAGGCTGCATGATGGG 768

|||||
DB 481 GCAATGTCTACATTAACCAATGACCCCAATCTGTCTCTGACAGGCTGCATGATGGG 540
QY 769 TTCAGGTTGTTAAAGCTAAATGATCATCCGCGCAAGTGCATGCTGTAATTAATTCACCA 828
|||||
DB 541 TTCAGGTTGTTAAAGCTAAATGATCATCCGCGCAAGTGCATGCTGTAATTAATTCACCA 600
QY 829 GGAATTAAGATGTAGTACACGCGGACCTTGATCGCATGAATAACAGTTGATCGTA 888
|||||
DB 601 GGAATTAAGATGTAGTACACGCGGACCTTGATCGCATGAATAACAGTTGATCGTA 660
QY 889 TGCAATATGGGCCAC-TCCACACAGAAATCGATGTGACCAAGCCTCC 934
|||||
DB 661 TGCAATATGGGCCACTTCCACACAGAAATCGATGTGACCAAGCCTTC 707
RESULT 11
BM837838
LOCUS
DEFINITION BM837838 691 bp mRNA linear EST 06-MAR-2002
K-EST0113987 S9SNU601 Homo sapiens cDNA clone S9SNU601-76-A11 5',
mRNA sequence.
ACCESSION BM837838
VERSION BM837838.1 GI:19194247
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 691)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 76 row: A column: 11
High quality sequence stop: 691.
FEATURES
source
1. 691
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="S9SNU601-76-A11"
/clone_1lb="S9SNU601"
/sex="M"
/tissue_type="Asclites"
/cell_type="Epithelial"
/cell_line="SNU-601"
/lab_host="Top10F"
/note="Organ: Stomach; Vector: PME18-FL3; Site 1: XhoI;
Site 2: XhoI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then deacapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including SfiI
site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized with Superscript II using SfiI
oligo-dT primer. After first strand synthesis, RNA was
degraded by NaOH treatment and cDNA was amplified by PCR
reaction. The PCR products were digested with SfiI and
cloned into DraIII- digested PME18S-FL3 vector. The
obtained cDNA vectors were used for transfection of
competent cells E. coli Top10F, by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."
BASE COUNT 175 a 151 c 194 g 171 t
ORIGIN
Query Match 53.1%; Score 687.8; DB 14; Length 691;

Best Local Similarity 99.78; Pred. No. 1.5e-199;
Matches 689; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 389 ATGATGGGGAGACTTAACCCACTGGGTTTATAGAGATATCCAAACGTTTAAAGA 448
    |||||||
Db 1 ATGATGGGGAGACTTAACCCACTGGGTTTATAGAGATATCCAAACGTTTAAAGA 60
QY 449 TCCGAGGCAATGTGGAGAGAGCGTGTGTTTCAAGGCTGTATCAGCTCTCCAAAG 508
    |||||||
Db 61 TCCGAGGCAATGTGGAGAGAGCGTGTGTTTCAAGGCTGTATCAGCTCTCCAAAG 120
QY 509 CTGGGAAGCTCTGTGTCCGGCCATGAACGTCATGATTTCTGTACCAACAGAGTTTG 568
    |||||||
Db 121 CTGGGAAGCTCTGTGTCCGGCCATGAACGTCATGATTTCTGTACCAACAGAGTTTG 180
QY 569 ATAACCTGTACTGCTGCCAGAGATCCATTTGGATGGCTGAAGAGACACAGATGTGA 628
    |||||||
Db 181 ATAACCTGTACTGCTGCCAGAGATCCATTTGGATGGCTGAAGAGACACAGATGTGA 240
QY 629 TGTGTGGTGGGAACAAGTGGTGTGTGTGGCTATAGGTAGGCAAGGGCTGTGTG 688
    |||||||
Db 241 TGTGTGGTGGGAACAAGTGGTGTGTGTGGCTATAGGTAGGCAAGGGCTGTGTG 300
QY 689 CTGCTCTCAAGCTCTGTGGAGCAATTTGTCTACATTAACGAAATCGACCCATCTGCTC 748
    |||||||
Db 301 CTGCTCTCAAGCTCTGTGGAGCAATTTGTCTACATTAACGAAATCGACCCATCTGCTC 360
QY 749 TGCAGGCTGCATGATGGGTTCAAGGTGTAAAGCTAAATGAAGTATCCGGCAAGTCG 808
    |||||||
Db 361 TGCAGGCTGCATGATGGGTTCAAGGTGTAAAGCTAAATGAAGTATCCGGCAAGTCG 420
QY 809 ATGTCGTAATACTGTGACAGGAATAAGATGTAGTACACAGGAGCATTGGATCGCA 868
    |||||||
Db 421 ATGTCGCAATACTGTGACAGGAATAAGATGTAGTACACAGGAGCATTGGATCGCA 480
QY 869 TGAATAACAGTGTGTATGATATGCAATATGGGCCACTCCACACAGAAATCGATGTGACA 928
    |||||||
Db 481 TGAATAACAGTGTGTATGATATGCAATATGGGCCACTCCACACAGAAATCGATGTGACA 540
QY 929 GCCTCCGCACTCCGAGCTGACGTGGAGCGAGTCTCTCAGGTGAGGACATGTATCT 988
    |||||||
Db 541 GCCTCCGCACTCCGAGCTGACGTGGAGCGAGTCTCTCAGGTGAGGACATGTATCT 600
QY 989 GGCAGATGGCAAGAGTGTCTCTCTGCGAGAGGCTCTACTCAATTGAGCTGCT 1048
    |||||||
Db 601 GGCAGATGGCAAGAGTGTCTCTCTGCGAGAGGCTCTACTCAATTGAGCTGCT 660
QY 1049 CCACAGTCCACCTTGTCTGTCCATCAC 1079
    |||||||
Db 661 CCACAGTCCACCTTGTCTGTCCATCAC 691
```

RESULT 12
BM454352
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BM454352 1098 bp mRNA linear EST 05-FEB-2002
AGENCOURT_6402556 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5497482
5' mRNA sequence.
BM454352
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 1098)
NIH-MGC http://mhc.ncl.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lou Staudt
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLNL12127 row: k column: 19
high quality sequence stop: 518.
Location/Qualifiers
1. 1098
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5497482"
/clone_1b="NIH_MGC_85"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 226 a 344 c 272 g 256 t
ORIGIN

Query Match 51.98; Score 673; DB 13; Length 1098;
Best Local Similarity 99.38; Pred. No. 7.7e-195;
Matches 676; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
QY 263 TAGCTGACGACCTGGCTGAGCTGAGTGTGACATGATGGGTGGAGGCCAAGTGA 322
    |||||||
Db 1 TAGCTGACGACCTGGCTGAGCTGAGTGTGACATGATGGGTGGAGGCCAAGTGA 60
QY 323 ATGACTTCTGTGTATGACCGCTGTGTGAACATGATGGGTGGAGGCCAAGTGA 382
    |||||||
Db 61 ATGACTTCTGTGTATGACCGCTGTGTGAACATGATGGGTGGAGGCCAAGTGA 120
QY 383 TCCTGGATGATGGGGAGACTTAACCCACTGGGTTTATAGAGATATCCAAACGTTT 442
    |||||||
Db 121 TCCTGGATGATGGGGAGACTTAACCCACTGGGTTTATAGAGATATCCAAACGTTT 180
QY 443 AGAAGATCCGAGGCAATGTGGAGAGAGCGTGTGTTTACAGGCTGTATCAGCTCT 502
    |||||||
Db 181 AGAAGATCCGAGGCAATGTGGAGAGAGCGTGTGTTTACAGGCTGTATCAGCTCT 240
QY 503 CCAAGCTGGGAAGCTGTGTCCGGCCATGAACGTCATGATTTGATGGCTGAAGAGCA 562
    |||||||
Db 241 CCAAGCTGGGAAGCTGTGTCCGGCCATGAACGTCATGATTTGATGGCTGAAGAGCA 300
QY 563 AGTTGATACTTGTCTGCTCCGAGATCCATTTGATGGCTGAAGAGCAACAG 622
    |||||||
Db 301 AGTTGATACTTGTCTGCTCCGAGATCCATTTGATGGCTGAAGAGCAACAG 360
QY 623 ATGTGATGTTGGTGGGAACAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 682
    |||||||
Db 361 ATGTGATGTTGGTGGGAACAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
QY 683 GCTGTGCTGCTCTCAAGCTCTTGAGCAATGTGTATACATTAACGAATCGACCCATCT 742
    |||||||
Db 421 GCTGTGCTGCTCTCAAGCTCTTGAGCAATGTGTATACATTAACGAATCGACCCATCT 480
QY 743 GTGCTGTGAGGCTGTGATGGGTTCAGGGTGAAGCTAAATGAAGTCAATCCGGC 802
    |||||||
Db 481 GTGCTGTGAGGCTGTGATGGGTTCAGGGTGAAGCTAAATGAAGTCAATCCGGC 540
QY 803 AAGTCGATGTCGTAATACTTGACAGGAATAAGATGTAGTGACACGGGAGCACTTG 862
    |||||||
Db 541 AAGTCGATGTCGTAATACTTGACAGGAATAAGATGTAGTGACACGGGAGCACTTG 600
QY 863 ATGCGATGAAAAACAGTGTATGATGCAATATGGGCCACTCCAAACAGAAATCGATG 922
    |||||||
Db 601 ATGCGATGAAAAACAGTGTATGATGCAATATGGGCCACTCCAAACAGAAATCGATG 660
QY 923 TGACCAAGCTCCGCACTCCGG 943
    |||||||
Db 661 TGACCAAGCTCCGCACTCCGG 681
```

```

RESULT 13
LOCUS      BF792750
DEFINITION BF792750      956 bp      mRNA      linear      EST 12-JAN-2001
            602253481F1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4345629 5',
            mRNA sequence.
ACCESSION  BF792750
VERSION     BF792750.1  GI:12097804
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 956)
            NIH-MGC http://mgs.ncl.nih.gov/.
AUTHORS     National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE       Unpublished (1999)
JOURNAL     Contact: Robert Strausberg, Ph.D.
            Email: cga@b-remail.nih.gov
COMMENT     Tissue Procurement: ATCC
            CDNA Library Preparation: Life Technologies, Inc.
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.llnl.gov
            Plate: L1AM9966 row: a column: 22
            High quality sequence stop: 671.
FEATURES
    source
        1..956
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:4345629"
        /clone_1lb="NIH_MGC_84"
        /tissue_type="adrenal cortex carcinoma, cell line"
        /lab_host="DH10B (phage-resistant)"
        /note="Organ: adrenal gland; Vector: pCMV-SPORT6; Site: 1:
        NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT
        primed. Average insert size 1.229 kb. Library enriched for
        full-length clones and constructed by Life Technologies.
        Note: this is a NIH_MGC library."
BASE COUNT  234 a      236 c      284 g      202 t
ORIGIN
Query Match      51.9%; Score 672.4; DB 12; Length 956;
Best Local Similarity 93.1%; Pred. No. 1e-194;
Matches 758; Conservative 0; Mismatches 51; Indels 5; Gaps 5;
QY 97 ATTTCACTCAGGAACGCTGCTCAGGGGAGAGAGCCCTGGCTGCTAAATAGTGGC 156
Db 1 ATTTCACTCAGGAACGCTGCTCAGGGGAGAGAGCCCTGGCTGCTAAATAGTGGC 60
QY 157 TGTACACATCATCAGCCAGACAGCGGTGTGATTGAGACACTCTGTCCTGGGGCT 216
Db 61 TGTACACATCATCAGCCAGACAGCGGTGTGATTGAGACACTCTGTCCTGGGGCT 120
QY 217 CAGTCCGCTGCTGCTGTGAACATCTCACTCAACTCAGATGAAGTAGCTGCAGCACTG 276
Db 121 CAGTCCGCTGCTGCTGTGAACATCTCACTCAACTCAGATGAAGTAGCTGCAGCACTG 180
QY 277 GCTGAGCTGGAAGTGCAGTGTTCGCTTGAAGGGCGAGTCAAGATGACTTCTGTGG 336
Db 181 GCTGAGCTGGAAGTGCAGTGTTCGCTTGAAGGGCGAGTCAAGATGACTTCTGTGG 240
QY 337 TGTATTGACCGCTGTGTGAACATGATGGGTGGCAGGCCAATGATCTGGATGATGG 396
Db 241 TGTATTGACCGCTGTGTGAACATGATGGGTGGCAGGCCAATGATCTGGATGATGG 300
QY 397 GGAGACTTAACCCACTGGGTTTATAAGATATCCAAAGCTGTTTAAGAGATCCGAGGC 456
Db 301 GGAGACTTAACCCACTGGGTTTATAAGATATCCAAAGCTGTTTAAGAGATCCGAGGC 360

```

```

QY 457 ATTTGGAAGAGAGCGCTGACTGCTGTTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAG 516
Db 361 ATTTGGAAGAGAGCGCTGACTGCTGTTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAG 420
QY 517 CTCTGTTTCCGGCCATGAACGCTCAATGATTTCTGTACCAAGAGAGTTGATTAAGTTG 576
Db 421 CTCTGTTTCCGGCCATGAACGCTCAATGATTTCTGTACCAAGAGAGTTGATTAAGTTG 480
QY 577 TACTGTCGCCGAGAATCCATTTTGGATGGCTGAAGAGACACAGATGTGATGTTGGT 636
Db 481 TACTGTCGCCGAGAATCCATTTTGGATGGCTGAAGAGACACAGATGTGATGTTGGT 540
QY 637 GGGAAA-CAAGTGTGTGTGTGCTATGTTGAGTA-GGCAAGGGCTGCTGTGCTGCTC 694
Db 541 GGGAAACCAAGTGTGTGTGCTATGTTGAGTAAGGTAAGGCAAGGGCTGCTGTGCTGCTC 600
QY 695 TCAAAGCTCTTGAGCAATGCTTACATTCAGCAATCGAACCCCACTGCTGCTGAGG 754
Db 601 TCAAAGCTCTTGAGCAATGCTTACATTCAGCAATCGAACCCCACTGCTGCTGAGG 658
QY 755 CCTGCATGATGGGTTCAGGGT-GGTAAAGCTAAATGAAGTCAATCGGCAAGTGCATGTC 813
Db 659 CCTGCATGATGGGTTCAGGGT-GGTAAAGCTAAATGAAGTCAATCGGCAAGTGCATGTC 718
QY 814 GTAATACTTGCACAGAAATAGAATAGTAGTACACAGGAGACACTTGGATCGCATGAA 873
Db 719 GTAATACTTGCACAGAAATAGAATAGTAGTACACAGGAGACACTTGGATCGCATGAA 778
QY 874 AACAGTGTATCGTATGCAATATGGGCCACTCCA 907
Db 779 AAACGTGTTTGCACCAATATGGGGCCACTCCA 812

RESULT 14
LOCUS      AL527928      938 bp      mRNA      linear      EST 13-FEB-2001
DEFINITION AL527928 LTI_NFL003_NBC3 Homo sapiens cDNA clone CS0DC027YC19 5
            prime, mRNA sequence.
ACCESSION  AL527928
VERSION     AL527928.1  GI:12791421
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 938)
            Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
            Full-length cDNA libraries and normalization
            Unpublished (2001)
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 Evry cedex - France
            Email: segref@genoscope.cns.fr, web : www.genoscope.cns.fr.
FEATURES
    source
        1..938
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="CS0DC027YC19"
        /clone_1lb="LTI_NFL003_NBC3"
        /sex="male"
        /tissue_type="neuroblastoma cells"
        /lab_host="DH10B"
        /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
        was primed with a NotI-oligo(dT) primer. Five prime end
        enriched, double-stranded cDNA was digested with Not I and
        cloned into the Not I and Eco RV sites of the pCMVSPORT 6
        vector. Library was normalized. Contact : Feng Liang Life
        Technologies, a division of Invitrogen 9800 Medical Center
        Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
        8371 Email : fliang@lifetech.com URL :
        http://fulllength.invitrogen.com"
BASE COUNT  241 a      208 c      275 g      212 t      2 others

```

ORIGIN

Query Match 51.4%; Score 666.2; DB 9; Length 938;
Best Local Similarity 99.6%; Pred. No. 8.3e-193;
Matches 676; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 1 TCCAAAGGCGACGACGACATTTCTGTGTGAAGACATCAAGCAGGAGATTTGGACGCCG 60
DB 261 TCCAAAGGCGACGACGACATTTCTGTGTGAAGACATCAAGCAGGAGATTTGGACGCCG 320
QY 61 GAGATTGAGATTGACAGACGACATGCTGCTGTGATTTCACTCAGGAAACGTCGTCAG 120
DB 321 GAGATTGAGATTGACAGACGACATGCTGCTGTGATTTCACTCAGGAAACGTCGTCAG 380
QY 121 GGGGAGAAAGCCCTTGGCTGTGCTAAATAGTGGCTGTACACACATCACAGCCAGACA 180
DB 381 GGGGAGAAAGCCCTTGGCTGTGCTAAATAGTGGCTGTACACACATCACAGCCAGACA 440
QY 181 GCGGTGTGATTGAGACACTCTGCTGCGCTGGGGCTCAGTCCGCTGCTGCTGTGTAAC 240
DB 441 GCGGTGTGATTGAGACACTCTGCTGCGCTGGGGCTCAGTCCGCTGCTGCTGTGTAAC 500
QY 241 ATCTACTCAACTCAGATGAGTGTGCTGACGACTGCTGAGGCTGAGTGTGACGTTC 300
DB 501 ATCTACTCAACTCAGATGAGTGTGCTGACGACTGCTGAGGCTGAGTGTGACGTTC 560
QY 301 GCTTGAAGGGGCGAGTCAAGATGACTTCTGTGTGTATGACCGCTGTGTGAACATG 360
DB 561 GCTTGAAGGGGCGAGTCAAGATGACTTCTGTGTGTATGACCGCTGTGTGAACATG 620
QY 361 GATGGGTGGCAGGCGCAACATGATCTGTGATGATGGGGAGACTTAACCCACTGGGTTAT 420
DB 621 GATGGGTGGCAGGCGCAACATGATCTGTGATGATGGGGAGACTTAACCCACTGGGTTAT 680
QY 421 AAGAATATCCAAACGTTTAAAGATCCGAGCATGTGGAAGAGAGCGGTGACTGT 480
DB 681 AAGAATATCCAAACGTTTAAAGATCCGAGCATGTGGAAGAGAGCGGTGACTGT 740
QY 481 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTCTCCGCCATGAACGTC 540
DB 741 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTCTCCGCCATGAACGTC 800
QY 541 AATGATTTCTGTACCAACAGATTGATACTTGTACTGCTGCCGAGATCCATTTTG 600
DB 801 AATGATTTCTGTACCAACAGATTGATACTTGTACTGCTGCCGAGATCCATTTTG 860
QY 601 GATGGCTGAAGAGGACACAGATGTGATGTTGGTGGAAACAAGTGGTGTGTGGC 660
DB 861 GATGGCTGAAGAGGACACAGATGTGATGTTGGTGGG-AAAGAGTGGTGTGTGGC 919
QY 661 TATGCTGAGGTAGGCAAG 679
DB 920 TATGCTGAGGTAGGCAAG 938

RESULT 15

LOCUS BG751929 930 bp mRNA linear EST 15-MAY-2001
DEFINITION 602731492F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4874951 5',
mRNA sequence.

ACCESSION BG751929
VERSION BG751929.1 GI:14062582
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 930)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov

Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1755 row: h column: 24
High quality sequence stop: 753.

FEATURES

source
1. 930
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4874951"
/clone_lib="NIH_MGC_43"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using Zap-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. 1"

BASE COUNT 261 a 210 c 237 g 222 t
ORIGIN

Query Match 51.1%; Score 661.8; DB 12; Length 930;
Best Local Similarity 95.6%; Pred. No. 1.8e-191;
Matches 756; Conservative 0; Mismatches 27; Indels 8; Gaps 7;

QY 512 GGAAGCTGTGTGTCGGCCATGAAGCTCAATGATCTGTACCAACAGAGTTGATA 571
DB 1 GGAAGCTGTGTGTCGGCCATGAAGCTCAATGATCTGTACCAACAGAGTTGATA 60
QY 572 ACTTGTACTGCTGCCGAGATCCATTTGGATGGCCTGAAGAGGACACAGATGTGATGT 631
DB 61 ACTTGTACTGCTGCCGAGATCCATTTGGATGGCCTGAAGAGGACACAGATGTGATGT 120
QY 632 TTGGTGGAAACAAGTGTGTGTGCTATGTGTAGGTAGGCAAGGCTGTGTGCTG 691
DB 121 TTGGTGGAAACAAGTGTGTGTGCTATGTGTAGGTAGGCAAGGCTGTGTGCTG 180
QY 692 CTCTCAAGCTCTTGAGCAATGTCTACATTAACGAATCGACCCCATCTGTCTCTGC 751
DB 181 CTCTCAAGCTCTTGAGCAATGTCTACATTAACGAATCGACCCCATCTGTCTCTGC 240
QY 752 AGGCTGATGATGGGTTCAGGGGTGTAAGCTAAATGAAGTCAATCCGCAAGTCGATG 811
DB 241 AGGCTGATGATGGGTTCAGGGGTGTAAGCTAAATGAAGTCAATCCGCAAGTCGATG 300
QY 812 TCGTAATACTTGACAGGAA-AATAAGATGTAGTGAACAGGGGACACTTGATCGCATG 870
DB 301 TCGTAATACTTGACAGGAAATAAGATGTAGTGAACAGGGGACACTTGATCGCATG 360
QY 871 AAAAAGCTTGTATCGTATGCAATATGGGCCACTCCACACAGAAATCGATGTGACACG 930
DB 361 AAAAAGCTTGTATCGTATGCAATATGGGCCACTCCACACAGAAATCGATGTGACACG 420
QY 931 CTCGGCACTCCGGAGCTGACGTGGGACGAGTACGTTCTCAGGTGAGACCATGTCTG 990
DB 421 CTCGGCACTCCGGAGCTGACGTGGGACGAGTACGTTCTCAGGTGAGACCATGTCTG 480
QY 991 CCAGATGGCAAAAGAGTGTCTCTCTGG-CAGAGGCTGCTTACTCAATTTGAGCTGCTC 1049
DB 481 CCAGATGGCAAAAGAGTGTCTCTCTGGAGAGGCTGCTTACTCAATTTGAGCTGATC 540
QY 1050 CACAGTCCACACT-TTGTCTGTGCTACACAGCCACACACAGAGCTTTGGCACTGATAG 1108
DB 541 CACAGTCCACACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
QY 1109 AACTCTAATAATGACCCGAGGGCGATACAAAGCAGATGTGTACTTGC-TTCCTAAGAAA 1167

Db 601 AACTCTATAATGCACCCGAGGGCGATACAGCAGATGTGTACTTGCGTTCCTAAGAAA 660
OY 1168 ATGGATGAATACGTG--CCAGCTTGCACTCTGCCAT-CATTGATGCCACCTTACAGAG 1224
Db 661 ATGGATGAATACGTGGCGCAGCTTTGCATCTGCCATCCATGTGATGGCCACCTTACAGAG 720
OY 1225 CTGACAGATGACCAAGCAAAATATCTGGGACTCAACAAAAATGGCCATTCAAAACCTAAT 1284
Db 721 CTGACAGATGACCAAGCAAAATATCTGGGACTCAAGGAACACTGGG-CATTCAAAACCTACT 779
OY 1285 TATTACAGATA 1295
Db 780 TATTACAGATA 790

Search completed: April 21, 2003, 03:59:07
Job time : 1760.62 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2003, 22:06:49 ; Search time 4760.08 Seconds

(without alignments)
11274.088 Million cell updates/sec

Title: US-09-782-051-1-COPY_1_1844

Sequence: 1 ggcgcggcgagtcgagct.....aacctattattacagatcac 1844

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenBank:

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_cm:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vl:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_or:*
- 21: em_ov:*
- 22: em_pat:*
- 23: em_ph:*
- 24: em_pl:*
- 25: em_ro:*
- 26: em_sts:*
- 27: em_sy:*
- 28: em_un:*
- 29: em_vl:*
- 30: em_htg_hum:*
- 31: em_htg_in:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rod:*
- 36: em_htg_man:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1844	100.0	2563	6	AX029176	AX029176 Sequence
2	1844	100.0	2677	9	AF315687	AF315687 Homo sapi
3	1840.8	99.8	2583	9	BC010681	BC010681 Homo sapi
4	1835.4	99.5	2552	9	BC016942	BC016942 Homo sapi
5	1803.8	97.8	2526	9	BC007576	BC007576 Homo sapi
6	1547	83.9	2258	9	HSU82761	U82761 Homo sapien
7	1503.8	81.6	1772	10	BC018218	BC018218 Mus muscu
8	1469.4	79.7	2510	9	HSM800298	AL049954 Homo sapi
9	999.2	54.2	2030	9	BC008349	BC008349 Homo sapi
10	999.2	54.2	5052	9	BC024325	BC024325 Homo sapi
11	995.6	54.0	5025	9	AB020635	AB020635 Homo sapi
12	979.8	53.1	1987	9	AK025372	AK025372 Homo sapi
13	651.8	35.3	1792	3	AY113501	AY113501 Drosophill
14	644	34.9	1369	9	AF035319	AF035319 Homo sapi
15	599	32.5	1323	9	BC003631	BC003631 Homo sapi
16	486.2	26.4	40320	2	AC015387	AC015387 Drosophill
17	486.2	26.4	172372	3	AC010110	AC010110 Drosophill
18	421.6	22.9	52423	2	AC020395	AC020395 Drosophill
19	421.6	22.9	80423	3	DROABDB	L07835 Drosophilla
20	421.6	22.9	175335	3	AC091636	AC091636 Drosophill
21	421.6	22.9	223098	3	AE003715	AE003715 Drosophill
22	421.6	22.9	338234	3	DMU31961	U31961 Drosophilla
23	412.6	22.4	2057	10	MUSSAHH	L32836 Mus musculu
24	411.6	22.3	5050	3	DMBX200	X13168 Drosophilla
25	411	22.3	2067	10	BC015304	BC015304 Mus muscu
26	411	22.3	190612	2	AC129093	AC129093 Mus muscu
27	406	22.0	1302	6	AX458351	AX458351 Sequence
28	402.4	21.8	2084	9	HUMAHCY2	M61832 Human S-ade
29	402.4	21.8	2175	9	BC010018	BC010018 Homo sapi
30	402.4	21.8	2183	9	BC011606	BC011606 Homo sapi
31	402.4	21.8	2355	9	AK097610	AK097610 Homo sapi
32	401.4	21.8	2029	10	RATFAHHA	M15185 Rat S-adeno
33	399	21.6	2088	5	XLA7835	AJ007835 Xenopus L
34	394.2	21.4	2129	5	XELAAH	L35559 Xenopus lae
35	375	20.3	1603	3	AY102668	AY102668 Drosophill
36	373.8	20.3	1564	3	DMAHCYGEN	X95636 D.melanogas
37	359	19.5	1617	3	AF080546	AF080546 Anopheles
38	358.8	19.5	81378	9	AL772411	AL772411 Human DNA
39	358.8	19.5	183584	30	AL360171	AL360171 Human DNA
40	354	19.2	1631	3	LEISADSH	M76556 L. donovani
41	346	18.8	1315	3	DDIAHHA	M19937 Slime mold
42	337	18.3	384	6	AX407699	AX407699 Sequence
43	334.4	18.1	173839	2	AC118421	AC118421 Rattus no
44	322.4	17.5	37187	8	AB004537	AB004537 Schizosac
45	322.4	17.5	43757	8	SPBC8D2	AL022072 S.pombe c

ALIGNMENTS

RESULT 1	AX029176	2563 bp	DNA	linear	PAT 16-SEP-2000
LOCUS	AX029176				
DEFINITION	Sequence 1 from Patent WO9814562.				
ACCESSION	AX029176				
VERSION	AX029176.1	GI:10190060			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
TITLE	1 (bases 1 to 2563)				
	Hart,D.N.				
	Enzyme having s-adenosyl-L-homocysteine hydrolase (ahcy) type				
	activity				

JOURNAL Patent: WO 9814562-A.1 09-APR-1998;
HART DEREK NIGEL JOHN (NZ)
FEATURES Location/Qualifiers
Source 1.2563
/organism="Homo sapiens"
/db_xref="taxon:9606"
<3.1847

CDS

/note="Open reading frame extends without a stop codon for the full 5' nucleotide sequence. The initiation codon has yet to be identified."

/codon_start=1

/protein_id="CAC09285.1"

/db_xref="GI:10190061"

/translation="RGVAGARSCFWFSCGHRCPAALGCRDAMATAPDKPTOLDA
GAGRRVDRVSEGAARAGRAPERGERGGGSSAAGRAGGMSMPDAMPLEGVSEELKO
AKEIEDAEKYSFMAVTYKAPKKQIOFADMOEFTKEPTKGRSLRSISOSTDSYS
SAASYTSSDDEYSPREKQOTNSKSSNECVKNIKQAEFGRREIEIAEDMSALISLR
KRAOGEKPLAGAKIVGCTHITATVLIETICALGACRMSACNISTONEVAAALAE
AGVAFAMKGESEDDFWCIDRCVNMDSQANMLDDGDLTHWYKKYPNVEKKIRG
IVEBSVTGVHRLYOLSKAGKLCVPANVNDSTKOKFDNLCCRESILDLKRTIDVM
FGKQVNVCGYGEVKGCCALKALGAIYITEIDPICALQACMDGFRVRLNEVIRQ
VDVITCTGNKNVTVREHLDRMKNSCIVCNMGHSNTEIDVTSRTPELTWERVRSQVD
HVIMPDKRVLAEGRLLNLSCTVPTFVLSITATQALALIELYNABEGRYKKQDQV
LLPKMDIYASLHLPFSDAHLTELTDQAKYLGLNKNGPFPKPNYRY"

BASE COUNT 646 a 604 c 677 g 636 t
ORIGIN

Query Match 100.0%; Score 1844; DB 6; Length 2563;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1844; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GGGCGGGGAGGTGAGAGTGGAGAGTGGTCTGTCTGTCTGTGGCCACCGTGGCT 60
Db 1 GGGCGGGGAGGTGAGAGTGGAGAGTGGTCTGTCTGTCTGTGGCCACCGTGGCT 60
QY 61 GTCCGGCTGCTTGGGCTGCCAACAAGAGCGGTGGCCACAGCACCCTCAGAACCGCA 120
Db 61 GTCCGGCTGCTTGGGCTGCCAACAAGAGCGGTGGCCACAGCACCCTCAGAACCGCA 120
QY 121 CGCAGCTCGACGAGGCGCGCCGAGAGGGGTGGCGGATCGCGTGTCCGAGGGCGCCGCG 180
Db 121 CGCAGCTCGACGAGGCGCGCCGAGAGGGGTGGCGGATCGCGTGTCCGAGGGCGCCGCG 180
QY 181 GGGCAGGGCGGGCGCCGAGAGGGGAAAGAGCGGGGGCGGGTCCAGCCGCTGGCC 240
Db 181 GGGCAGGGCGGGCGCCGAGAGGGGAAAGAGCGGGGGCGGGTCCAGCCGCTGGCC 240
QY 241 GGGCGGGCGGGGAGATGTCATGCTGACGCGATGCGGTGCGGGGGTCCGGAGAGCG 300
Db 241 GGGCGGGCGGGGAGATGTCATGCTGACGCGATGCGGTGCGGGGGTCCGGAGAGCG 300
QY 301 TGAAGCAGGCGCAGAGAGAGCGCGGAGAGTACTCTTCATGCGCACCGTCAACA 360
Db 301 TGAAGCAGGCGCAGAGAGAGCGCGGAGAGTACTCTTCATGCGCACCGTCAACA 360
QY 361 AGCGCGCCCAAGAGCAATTCAGTTTGTGATGACATGACAGAGTTTCAACCAATTCCCA 420
Db 361 AGCGCGCCCAAGAGCAATTCAGTTTGTGATGACATGACAGAGTTTCAACCAATTCCCA 420
QY 421 CCAAAACTGGCCGAAGATTTTGTCTGCTGATCTCACAGTCTCCACTGACAGCTACA 480
Db 421 CCAAAACTGGCCGAAGATTTTGTCTGCTGATCTCACAGTCTCCACTGACAGCTACA 480
QY 481 GTTCAGCTGCATCTACACAGATAGCTGTGATGAGGTTTCCCGAGAGAGAGCAGC 540
Db 481 GTTCAGCTGCATCTACACAGATAGCTGTGATGAGGTTTCCCGAGAGAGAGCAGC 540
QY 541 AAACCAACTCCAAGGCGCAGCAATTTCTGTGTAAGAACATCAAGCAGAGCAATTTG 600
Db 541 AAACCAACTCCAAGGCGCAGCAATTTCTGTGTAAGAACATCAAGCAGAGCAATTTG 600
QY 601 GACGCGGGAGATGAGATGACAGCAAGACATGTCTGCTGATTTCACTCAGAAAC 660
Db 601 GACGCGGGAGATGAGATGACAGCAAGACATGTCTGCTGATTTCACTCAGAAAC 660
```

```
Db 601 GACGCGGGAGATGAGATGACAGCAAGACATGTCTGCTGATTTCACTCAGAAAC 660
QY 661 GTGCTCAGGGGGAGAGCCCTTGCTGTGCTAAATAGTGGCTGTACACACATCACAG 720
Db 661 GTGCTCAGGGGGAGAGCCCTTGCTGTGCTAAATAGTGGCTGTACACACATCACAG 720
QY 721 CCCAGACGCGGTGTGATGAGACACTCTGTGCCCCGTGGGGCTCAGTGCCGCTGTG 780
Db 721 CCCAGACGCGGTGTGATGAGACACTCTGTGCCCCGTGGGGCTCAGTGCCGCTGTG 780
QY 781 CTTGTACATCTACTCACTCAGATGAGTAGCTGCAGCAGCTGAGGCTGAGGTTG 840
Db 781 CTTGTACATCTACTCACTCAGATGAGTAGCTGCAGCAGCTGAGGCTGAGGTTG 840
QY 841 CAGTGTGCTTGGAGAGGGCGAGTCAGAGATGACTTCTGTGTGTATTGACCGCTGTG 900
Db 841 CAGTGTGCTTGGAGAGGGCGAGTCAGAGATGACTTCTGTGTGTATTGACCGCTGTG 900
QY 901 TGAACATGATGGGTGGCAGGCGCAACATGATCCTGGATGATGGGGAGACTTAACCACT 960
Db 901 TGAACATGATGGGTGGCAGGCGCAACATGATCCTGGATGATGGGGAGACTTAACCACT 960
QY 961 GGGTTATAGAAGTATCCAAACGTTTAAAGAGATCCGAGCATTTGGAAGAGAGCG 1020
Db 961 GGGTTATAGAAGTATCCAAACGTTTAAAGAGATCCGAGCATTTGGAAGAGAGCG 1020
QY 1021 TGACTGTGTTCAACAGCGCTGATACCTCTCCAAAGCTGGGAGCTGTGTCCGGCCA 1080
Db 1021 TGACTGTGTTCAACAGCGCTGATACCTCTCCAAAGCTGGGAGCTGTGTCCGGCCA 1080
QY 1081 TGAAGCTCATGATTTCTGTTACCAACAGAGATTGATAACTGTGCTGCGGAGAAAT 1140
Db 1081 TGAAGCTCATGATTTCTGTTACCAACAGAGATTGATAACTGTGCTGCGGAGAAAT 1140
QY 1141 CCATTTGGATGGCCTGAAGAGGACCAACAGATGTGATGTTGGTGGAAACAAGTGTG 1200
Db 1141 CCATTTGGATGGCCTGAAGAGGACCAACAGATGTGATGTTGGTGGAAACAAGTGTG 1200
QY 1201 TGTGTGCTATGTTGAGTAGAGCAAGGCTGTGCTGTCTCAAAAGCTCTTGAGCAAA 1260
Db 1201 TGTGTGCTATGTTGAGTAGAGCAAGGCTGTGCTGTCTCAAAAGCTCTTGAGCAAA 1260
QY 1261 TTGTCTACATTAACGAATGACCCCATCTGTGCTGTGACGCGCTCATGATGGTTCA 1320
Db 1261 TTGTCTACATTAACGAATGACCCCATCTGTGCTGTGACGCGCTCATGATGGTTCA 1320
QY 1321 GGGTGTAAAGCTAAATGAAGTCAATCCGGCAAGTGCATGCTGTAATACTTGACAGGAA 1380
Db 1321 GGGTGTAAAGCTAAATGAAGTCAATCCGGCAAGTGCATGCTGTAATACTTGACAGGAA 1380
QY 1381 ATAGAATGTAGTGACACGCGGACACTTGGATGCGATGAATAACAGTTGTATCGTATGCA 1440
Db 1381 ATAGAATGTAGTGACACGCGGACACTTGGATGCGATGAATAACAGTTGTATCGTATGCA 1440
QY 1441 ATATGGGCCACTCCAACACAGAAATCGATGTGACAGCGCTCCGCACTCCGAGCTGACGT 1500
Db 1441 ATATGGGCCACTCCAACACAGAAATCGATGTGACAGCGCTCCGCACTCCGAGCTGACGT 1500
QY 1501 GGAAGCAGTACGTTCTCAGGTGACCATGTCACTGTGCGCAGATGGCAACAGAGTTGTCC 1560
Db 1501 GGAAGCAGTACGTTCTCAGGTGACCATGTCACTGTGCGCAGATGGCAACAGAGTTGTCC 1560
QY 1561 TCCGTGCAAGAGGGTCTCTACTCAATTGAGCTGCTCCACAGTTCCCACTTGTCTGT 1620
Db 1561 TCCGTGCAAGAGGGTCTCTACTCAATTGAGCTGCTCCACAGTTCCCACTTGTCTGT 1620
QY 1621 CCATCAGACGCCACACACAGGCTTTGGCACTGATAGAACTCTAATATGACACCGAGGGGC 1680
Db 1621 CCATCAGACGCCACACACAGGCTTTGGCACTGATAGAACTCTAATATGACACCGAGGGGC 1680
QY 1681 GATACAGCAGGATGTGTACTGCTCTTAAGAAATGATGATGATGATGATGATGATGATGAT 1740
Db 1681 GATACAGCAGGATGTGTACTGCTCTTAAGAAATGATGATGATGATGATGATGATGATGAT 1740
```

[illegible]

RESULT 2	AF315687	LOCUS	AF315687	DEFINITION	Homo sapiens S-adenosylhomocysteine hydrolase-like protein mRNA, complete cds.	2677 bp	mRNA	linear	PRI 22-MAR-2002
ACCESSION	AF315687	VERSION	AF315687.1	GI:	16588686				

```
KEYWORDS      .
SOURCE        Homo sapiens,
ORGANISM      Homo sapiens
```

REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2677)
Dekker, J.W., Budhia, S., Angel, N.Z., Cooper, B.J., Clark, G.J.,
Hart, D.N. and Kato, M.

TITLE Identification of an S-adenosylhomocysteine hydrolase-like transcript induced during dendritic cell differentiation
JOURNAL Immunogenetics 53 (12), 993-1001 (2002)

MEDLINE	21901265
PUBMED	11904675

REFERENCE	2 (bases 1 to 2677)
AUTHORS	Dekker, J.W., Budhla, S., Angel, N.Z., Cooper, B.J., Clark, G.J.,

TITLE Direct Submission
JOURNAL Submitted (23-OCT-2000) Dendritic Cell Research, Mater Medical Research Institute, Level 3, Aubigny Place, South Brisbane, Queensland 4101, Australia

FEATURES	Location/Qualifiers
source	1..2677

```

/misc_feature
1. .114
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p21"
/clone="211(1)B".
/cell_line="Hodgkin's disease cell line L428"
1. .114

```

369.	1961	/note--"derived from 5' RACE"
CDS		

```

/note="DCAL; dendritic cell expressed AHcy-like protein;
expressed in dendritic cells in blood, langerhans cells
and dermal dendritic cells; mRNA increased during monocyte
differentiation to monocyte-derived dendritic cells in the
presence of GM-CSF and IL-4"
/codon_start=1
/product="S-adenosylhomocysteine hydrolase-like protein"
/protein_id="AAL26869.1"
/db_xref="GI:16588687"

```

/translation="MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMAVTYKAPKKOIQ
 ADDMQEFTKPTKTRSLRSISQSTDSYSSAASYETDSSDEVSPEKOOTNSKGS
 SNECVKNIKQAEFGREIEIAEDMSALISLRKAQOGEKPLAGAKIVGCTHITQAQV
 LIETLCALGAQCRWSACNIYSTQNEVAALAEAGVAFAFWKGESEDEDFWCIDRCVNM
 DGNQANMILDDGDLTHWVYKYPNFKIRGIVEESVTGVHRLYQSKAGLCPAMPAM
 NVNDSVTKQKFDNLCCRESILDLKRTTYDMFGGKQVVCGYGEVKGCCALKALG
 AIYITEIDPICALQACMDGFRVYKNEVIRQDVITCTGNKNVYREHLDRMKNSC
 IYCNMGHSNTEIDVTSLRPELWEYRYSQVDHVIWPDGKRVVLLAEGRLNLSCSTV
 PTFVLSTATITQALILIELYNAPEGRYKQDVYLLPKKMDYVASLHPSFDAHLTELT
 DQAKYLGINKNGPFKNPYRY"

BASE COUNT	657 a	648 c	708 g	664 t
ORIGIN				

Query Match	100.0%;	Score 1844;	DB 9;	Length 2677;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1844; Conservative	0;	Mismatches	0;	Indels 0;
				Gaps 0;

QY 1 GGGCGGGCAGGTCGGAGCTCGGAGCTCTGCTTCTCTGTGGCCACCGTGCT 60
|||||
Db 115 GGGCGGGCAGGTCGGAGCTCGGAGCTCTGCTTCTCTGTGGCCACCGTGCT 174
QY 61 GTCCGGCTGCCTTGGGCTGCCGAACAGCAGCGGTGGCCACAGCCTCAGAAGCCGA 120
|||||
Db 175 GTCCGGCTGCCTTGGGCTGCCGAACAGCAGCGGTGGCCACAGCCTCAGAAGCCGA 234

Db 115 GGCAGGGCAGGTCGAGCTCGAGCTGCTGCTTCTGCTTGTGGCCACCGTGGCT 174

Db 175 GTCGGCTGCCTTGGCTGCCGAACAGCAGCGTGGGCCACAGCACCCTCAGAAGCCGA 234

QY 121 CCGACCTCGACGCGAGGGCCCGGCAGAGAGGTGGCCGATCGCGTCTCGGAGGCGCCCGCC 180

Db 235 CGCAGCTGACGCGAGGGGCCGAGAGGTGGGCGATCGCGTGTGGAGGGCCGCCG 294

181 GGGCAGCGCGCGCGCGCGCCACAGAGGGGAAAGAGCGCGCGCGCGGTCAAGCCGCTGGCC 240

DB 295 GGGCAGGGCGGGCGGGCGGGCCAGAGGGGGAAAGAGGGCGGGCGGGCGGGGTCAAGCCGCTGGCC 354

355 GGGGCGGGGGGGGGAATGTCGATGCGCTGAGACCGATGCGCGTGGGGGTGGGGAGGAGC 414

301 TGAACGAGGCCAAGGAGATCGAGGACGCCGAGAACTACTCTTCATGCGCCACCGTCACCA 360

Db 415 TGAAGCAGGCCAAGAGATCGAGACGCCGAGAAGTACTCCTTCATGGCCACCGTCACCA 474

QY 361 AGGCGCCCAAGAAGCAATCCAGTTTGGCTGATGACATGCAGGAGTTTACCAAAATTCCCA 420

Db 475 AGGGCCCAAGAAGCAATCCAGTTTGTGATGACATGCAGAGTTCACCAATTCCCA 534

QY 421 CCAAACTGGCCGAAGATCTTTGTCTCGCTCGATCTCACAGTCCCTCCACTGACAGCTACA 480

535 CCAAAACTGGCCGAAGATCTTGTCTCGCTCGATCTCACAGTCTCTCCACTGACAGCTACA 594

[illegible]

541 AAACCAACTCCAGGGGAGCAGCAATTTTGTGTGTGAGAACATCAAGCAGCAGAAATTTG 600

Db 655 AACCACTCCAGGGCAGCAGCATTTCTGTGTGAAGACATCAAGCAGCAGCATTTTG 714

601 GACGCCGGAGATTGAGATTGCAGACAGACATGTCTGCTCTGATTTCACTCAGGAAC 660

Db 715 GACGCCGGAGATTGAGATTCGACAGCAGACATGTCCTCTGATTTCACTCAGGAAC 774

661 GTGCTCAGGGGGGAGAGCCCTTGGCTGTGCTAAATAGTGGGCTGTACACACATCACAG 720

Db 775 GTCCTCAGGGGGGAGAGCCCTTGGCTGGTGTCTAAATAGTGGGCTGTACACACATCACAG 834

CCAGACAGCGGTGTGATGAGACACTCTGTGCCCTGGGGCTCAGTCCGCTGGTCTG 780

[illegible]

Db. 895 CTTGTACATCTACTCACTCAGAAATGAACTGACGACACTGGCTGAGGCTGGAGTTG 954

841 CAGTGTTCGCTTGGAGGCGAGTCAGAGATGACTTCTGTGTGTATTTGACCGCTGTG 900

Db 955 CAGTGTCCCTTGCAGAGGCCAGACAGAGATGACTTCTGTGTGTGTATGACCGCTGTG 101

QY 901 TGAACATGGATGGGTGGCAGGCCAACATGATCCTGGATGATGGGGGAGACTTAACCACT 960

Db 1015 TGAACATGGATGGGTGGCAGGCCACATGATCCTGGATGATGGGGGAGACTTAACCCACT 107

QY 961 GCGTTATAAGAGTATCCAAACGCTTTAAGAAGATCCGAGCATTTGGAGAAGAGCG 102

[illegible][illegible]

```

QY 1081 TGAAGTCAATGATCTCTTACCAACAGAGTTGATACCTGTACTGCTGCCGAGAT 1140
    |||||||
DB 1195 TGAACGTCAATGATCTCTTACCAACAGAGTTGATACCTGTACTGCTGCCGAGAT 1254
QY 1141 CCATTTTGGATGGCCTGAAGAGGACACAGATGTGATGTTGGTGGGAACAAGTGTGG 1200
    |||||||
DB 1255 CCATTTTGGATGGCCTGAAGAGGACACAGATGTGATGTTGGTGGGAACAAGTGTGG 1314
QY 1201 TGTGTGGCTATGCTGAGGTAGGCAAGGGCTGCTGCTCTCTCAAGCTCTTGAGCAA 1260
    |||||||
DB 1315 TGTGTGGCTATGCTGAGGTAGGCAAGGGCTGCTGCTCTCTCAAGCTCTTGAGCAA 1374
QY 1261 TTGCTACATTAACGAAATCGACCCCATCTGCTCTGACAGCCCTGCATGGATGGTTCA 1320
    |||||||
DB 1375 TTGCTACATTAACGAAATCGACCCCATCTGCTCTGACAGCCCTGCATGGATGGTTCA 1434
QY 1321 GGGTGTAAAGCTAAATGAAGTCAATCCGGCAAGTGTGTAATAACTGACAGAGAA 1380
    |||||||
DB 1435 GGGTGTAAAGCTAAATGAAGTCAATCCGGCAAGTGTGTAATAACTGACAGAGAA 1494
QY 1381 ATAAGATGTAGTGACACGGGAGCAGCTGATCCGATGAAAAACAGTGTATCGTATGCA 1440
    |||||||
DB 1495 ATAAGATGTAGTGACACGGGAGCAGCTGATCCGATGAAAAACAGTGTATCGTATGCA 1554
QY 1441 ATATGGCCACTCCACACAGAAATCGATGTGACACAGCCCTCCGACACTCCGAGCTGACGT 1500
    |||||||
DB 1555 ATATGGCCACTCCACACAGAAATCGATGTGACACAGCCCTCCGACACTCCGAGCTGACGT 1614
QY 1501 GGGAGCGAGTACGTTCTCAGGTGAGCAGCATGTCTGCGCCAGATGGCAACAGAGTTGTC 1560
    |||||||
DB 1615 GGGAGCGAGTACGTTCTCAGGTGAGCAGCATGTCTGCGCCAGATGGCAACAGAGTTGTC 1674
QY 1561 TCCTGGCAGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTTCACCTTTGTTCTGT 1620
    |||||||
DB 1675 TCCTGGCAGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTTCACCTTTGTTCTGT 1734
QY 1621 CCATCACAGCCACACACAGGCTTTGGCAGCTGATGAACCTTATATGACACCCGAGGGGC 1680
    |||||||
DB 1735 CCATCACAGCCACACACAGGCTTTGGCAGCTGATGAACCTTATATGACACCCGAGGGGC 1794
QY 1681 GATACAAGCAGAGTGTGTACTGCTCTCTCTTAAGAAATGATGATAGCTGCGCAGCTTGC 1740
    |||||||
DB 1795 GATACAAGCAGAGTGTGTACTGCTCTCTCTTAAGAAATGATGATAGCTGCGCAGCTTGC 1854
QY 1741 ATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1800
    |||||||
DB 1855 ATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1914
QY 1801 GACTCAACAAAAATGGGCCATTCMAACCTTATATACAGATAC 1844
    |||||||
DB 1915 GACTCAACAAAAATGGGCCATTCMAACCTTATATACAGATAC 1958

```

```

RESULT 3
LOCUS BC010681 2583 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, S-adenosylhomocysteine hydrolase-like 1, clone
ACCESSION BC010681
VERSION BC010681.1 GI:14715037
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 2583)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.ncl.nih.gov

```

COMMENT

Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: villalob@bcm.tmc.edu
 Villalon, D.R., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
 A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
 Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAP Plate: 13 Row: B Column: 21
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA g1: 5729723.

FEATURES

source

1. 2583

/organism="Homo sapiens"

/db_xref="LocusID:10768"

/db_xref="taxon:9606"

/clone="MGC:8936 IMAGE:3853747"

/tissue_type="Colon, adenocarcinoma"

/clone_1db="NIH_MGC_65"

/lab_host="DH10B"

/note="Vector: PCMV-SPORT6"

367, 1869

CDS

/codon_start=1

/product="S-adenosylhomocysteine hydrolase-like 1"

/protein_id="AAH10681.1"

/db_xref="GI:14715038"

/translation="MATVTKAPKKQIQFADDMOEFKPTKGRRLSRSISOSTDS

YSSASYSDDSDDEVSPREKQNTNSKSSNFCVKNIKQAEFRREIEIAEODMSALIS

LKRRAQGEKPLAGAKIVGCTHTAQTAVLIETLCALGAQCRSACNISTONEVAAL

AEAGYAVFAWKGESEDDFWWCIDRCVNMDSVTKOKFDNLCCRESILDLKRTTD

RGIVESVTGVHRLYQLSKAGKLCVPMNVNDSVTKOKFDNLCCRESILDLKRTTD

VMFGKOVVVCYGEVGVKSCCALKALGAIYITEIDPICALQACMDGFRVKNLEVI

ROYDVITCTGNKNVTRHLDKMSNCIVCMGHSNTEIDVSLRPELTWERVRSQ

VDHYIPDGRKRVLLAEGRLNLSCTVPTFVLSITATQALALIELYNAPGRYKOD

VLLPKKMDDEVASLHLPFDALHLELTDQAKYLGINKNGEPKNYRY"

BASE COUNT

648 a 606 c 688 g 641 t

ORIGIN

Query Match 99.8%; Score 1840.8; DB 9; Length 2583;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1842; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 GGGCGGGCAGAGTGGAGCTGGAGCTGCTTCTGTTCTTGTGGCCACCGTGGCT 60
    |||||||
DB 23 GGGCGGGCAGAGTGGAGCTGGAGCTGCTTCTGTTCTTGTGGCCACCGTGGCT 82
QY 61 GTCCGGCTGCTTGGCTGCCGAACAGACAGAGCGGTGGCCACAGCAGCTCAGAGCCGA 120
    |||||||
DB 83 GTCCGGCTGCTTGGCTGCCGAACAGACAGAGCGGTGGCCACAGCAGCTCAGAGCCGA 142
QY 121 CGCAGCTGACGACGAGGGCCGCGAGAGGCTGGCGATCGCGTGTGAGAGGGCCGCGCC 180
    |||||||
DB 143 CGCAGCTGACGACGAGGGCCGCGAGAGGCTGGCGATCGCGTGTGAGAGGGCCGCGCC 202
QY 181 GGGCAGGGCGGGCGGCCAGAGGGGAAAGAGCGGGGGGGGGGGGTACAGCCGTGGCC 240
    |||||||
DB 203 GGGCAGGGCGGGCGGCCAGAGGGGAAAGAGCGGGGGGGGGGGGTACAGCCGTGGCC 262
QY 241 GGGCGGGCGGGGAGATGCTGATGCTGACGCGGATGCGGCTGCCGGGGTCCGGGAGAGC 300
    |||||||
DB 263 GGGCGGGCGGGGAGATGCTGATGCTGACGCGGATGCGGCTGCCGGGGTCCGGGAGAGC 322
QY 301 TGAAGCAGGCCAAGAGATCGAGGAGCGCCGAGAGTACTCTTCATGGCCACCGTCACGA 360
    |||||||
DB 323 TGAAGCAGGCCAAGAGATCGAGGAGCGCCGAGAGTACTCTTCATGGCCACCGTCACGA 382

```

QY	361	AGCGCCCAAGAACAAATCCAGTTTGCTGATGACATGCAGGAGTTCCACCAATTTCCCA	420
Db	383	AGCGCCCAAGAACAAATCCAGTTTGCTGATGACATGCAGGAGTTCCACCAATTTCCCA	442
QY	421	CCAAACTGGCCGAAGATCTTTGTCGCTCGATGCTCAGTCCCTCCACTGACAGCTACA	480
Db	443	CCAAACTGGCCGAAGATCTTTGTCGCTCGATGCTCAGTCCCTCCACTGACAGCTACA	502
QY	481	GTTACGCTGCATCCTACACAGATAGCTCTGATGATGAGGTTTCTCCCGAGAGAACGAC	540
Db	503	GTTACGCTGCATCCTACACAGATAGCTCTGATGATGAGGTTTCTCCCGAGAGAACGAC	562
QY	541	AAACCACTCCAAGGGCAGCAGCAATTTCTGTGTGAAGAACATCAAGCAGGAGAAATTG	600
Db	563	AAACCACTCCAAGGGCAGCAGCAATTTCTGTGTGAAGAACATCAAGCAGGAGAAATTG	622
QY	601	GACGCCGGAGATTGAGATTGCAGAGCAAGACATGTCCTCTGATTTCACTCAGGAAC	660
Db	623	GACGCCGGAGATTGAGATTGCAGAGCAAGACATGTCCTCTGATTTCACTCAGGAAC	682
QY	661	GTGCTCAGGGGAGAGACCCCTTGCTGTGCTAAATAGTGGGCTGTACACACATCAG	720
Db	683	GTGCTCAGGGGAGAGACCCCTTGCTGTGCTAAATAGTGGGCTGTACACACATCAG	742
QY	721	CCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGGGGGCTCAGTGCCGCTGTG	780
Db	743	CCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGGGGGCTCAGTGCCGCTGTG	802
QY	781	CTGTACATCTACTCAACTCAGATGAAGTAGCTGCAGCACTGGCTGAGGCTGAGATTG	840
Db	803	CTGTACATCTACTCAACTCAGATGAAGTAGCTGCAGCACTGGCTGAGGCTGAGATTG	862
QY	841	CAGTGTTCGCTTGAAGGGCGAGTCAGAAGATGACTTCTGGTGTATTGACCGCTGTG	900
Db	863	CAGTGTTCGCTTGAAGGGCGAGTCAGAAGATGACTTCTGGTGTATTGACCGCTGTG	922
QY	901	TGAACATGATGGGTGGCAGGCCCAACATGATCCTGATGATGGGGAGACTTAACCACT	960
Db	923	TGAACATGATGGGTGGCAGGCCCAACATGATCCTGATGATGGGGAGACTTAACCACT	982
QY	961	GGGTTTATAAGAGATATCCAAACGTGTTTAAGAAAGATCCGAGGCATTGTGGAAGAGCG	1020
Db	983	GGGTTTATAAGAGATATCCAAACGTGTTTAAGAAAGATCCGAGGCATTGTGGAAGAGCG	1042
QY	1021	TGACTGCTGTTCAAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTGTTCCGGCCA	1080
Db	1043	TGACTGCTGTTCAAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTGTTCCGGCCA	1102
QY	1081	TGAACGTCAATGATCTCTGTACCAACAGAAATTGATACTTGTACTGCTGCCGAGAAAT	1140
Db	1103	TGAACGTCAATGATCTCTGTACCAACAGAAATTGATACTTGTACTGCTGCCGAGAAAT	1162
QY	1141	CCATTTGGATGGCCTGAAGAGGACCACAGATGTGATGTTTGGTGGAACAAGTGGTGG	1200
Db	1163	CCATTTGGATGGCCTGAAGAGGACCACAGATGTGATGTTTGGTGGAACAAGTGGTGG	1222
QY	1201	TGTGTGGCTATGTTAGGTAGGCTAGGCAAGGGCTGCTGTGCTCTCAAAAGCTCTTGAGCAA	1260
Db	1223	TGTGTGGCTATGTTAGGTAGGCTAGGCAAGGGCTGCTGTGCTCTCAAAAGCTCTTGAGCAA	1282
QY	1261	TTGTCTACATTAACCGAAATCGACCCCATCTGTGCTCTGCAAGGCTGCATGATGGTTCA	1320
Db	1283	TTGTCTACATTAACCGAAATCGACCCCATCTGTGCTCTGCAAGGCTGCATGATGGTTCA	1342
QY	1321	GGGTGTAAGCTAAATGAAGTCATCCGGCAATTCGATGCTAATACTTGACAGGAA	1380
Db	1343	GGGTGTAAGCTAAATGAAGTCATCCGGCAATTCGATGCTAATACTTGACAGGAA	1402
QY	1381	ATAAGAAATGTAGTACACCGGGAGCACTTGGATCGCATGAAAAACAGTTGTATCGTATGCA	1440
Db	1403	ATAAGAAATGTAGTACACCGGGAGCACTTGGATCGCATGAAAAACAGTTGTATCGTATGCA	1462

QY	1441	ATATGGGCCACTCCAAACACAGAAATCGATGTGACCAGCCTCCGCACCTCCGGAGCTGACGT	1500
Db	1463	ATATGGGCCACTCCAAACACAGAAATCGATGTGACCAGCCTCCGCACCTCCGGAGCTGACGT	1522
QY	1501	GCGAGCGAGTACGTTCTCAGGTGCAACCATGTCTGTGCGCAGATGGCAAACGAGTTGTCC	1560
Db	1523	GCGAGCGAGTACGTTCTCAGGTGCAACCATGTCTGTGCGCAGATGGCAAACGAGTTGTCC	1582
QY	1561	TCCTGGCAGAGGGTGCTCTACTCAATTGTAGCTGCTCCACAGTTCACACCTTTGTCTGT	1620
Db	1583	TCCTGGCAGAGGGTGCTCTACTCAATTGTAGCTGCTCCACAGTTCACACCTTTGTCTGT	1642
QY	1621	CCATCACAGCCACAACACAGGCTTTGGCACTGATAGAAGCTTAATATGCACCCGAGGGGC	1680
Db	1643	CCATCACAGCCACAACACAGGCTTTGGCACTGATAGAAGCTTAATATGCACCCGAGGGGC	1702
QY	1681	GATACAAGCAGGATGTGTACTTGCTTCTTAAGAAAATGGATGAATACGTTGCCAGCTTGC	1740
Db	1703	GATACAAGCAGGATGTGTACTTGCTTCTTAAGAAAATGGATGAATACGTTGCCAGCTTGC	1762
QY	1741	ATCTGCCATCATTTGATGCCCCAAGCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG	1800
Db	1763	ATCTGCCATCATTTGATGCCCCAAGCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG	1822
QY	1801	GACTCAACAAAAATGGGCCATTCAACCTAATTATTACAGATAC	1844
Db	1823	GACTCAACAAAAATGGGCCATTCAACCTAATTATTACAGATAC	1866

RESULT 4	
BC016942	
LOCUS	2552 bp mRNA linear PRI 09-NOV-2001
DEFINITION	Homo sapiens, S-adenosylhomocysteine hydrolase-like 1, clone
ACCESSION	MGC:21453 IMAGE:3450568, mRNA, complete cds.
VERSION	BC016942
KEYWORDS	BC016942.1 GI:16877386
SOURCE	MGC.
ORGANISM	Homo sapiens.
	Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2552)
Strausberg, R.
Direct Submission
Submitted (05-NOV-2001) National Institutes of Health, Mammalian

REMARK	COMMENT
NIH-MGC Project URL: http://mgc.ncl.nih.gov	Contact: MGC help desk

cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 20 Row: a Column: 15
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

```
FEATURES
    source
        location/qualifiers
            1. .2552
                /organism="Homo sapiens"
                /db_xref="LocusID:10768"
                /db_xref="taxon:9606"
                /clone="MGC:21453 IMAGE:3450568"
```

Page 5

```

/tissue_type="Placenta, choriocarcinoma"
/clone_1lb="NIH_MGC_10"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
338. .1840

```

```

/product="S-adenosylhomocysteine hydrolase-like 1"
/protein_id="AAH16942.1"
/db_xref="GI:16877387"
/translation="MATVYTKAPKKOIQFADDMOEFTKPTKGRSLRSRISQSTDS
YSSASYSYTDSSDDEVSPREKQDNTSKGSSNFCVKNIKQAEFGRELEIAEQDMSALIS
LRRRAOGEKPLAGAKIVGCTHTTAQTAVLIETICALGAOCRMSACNIYSTONEVAAL
AAGVAVFAWKGESEDEDFWMCIDRCVNMDDGQAMNILLDDGDLTHWVYKKYPNVFKI
RGITYESVTGVHRLYQLSKAGKLCVPAAMVNDSVTKQKFDNLXCCRSLDGLKRTD
VMGGKQVNVCGYGEYVTEGKCCALALGATVITYEIDPICALQAMGDFRVVKLNEVI
ROYDAVITCTGNKNVYTRREHLDRKNSCIVCNMGHSTNTEIDVYSLRTPELTWERVRSQ
VDHYIWPDGKRVVLLAEGRLNLSCTVPTFVLSITATQALAILLYNAPEGRYKQD
VLLPKRDEYVASLHLPSEDAHLTELTDDQAKYLGLNKNGPEKPNYYRY"

```

BASE COUNT	644 a	600 c	671 g	637 t
ORIGIN				

Query Match	99.58;	Score 1835.4;	DB 9;	Length 2552;
Best Local Similarity	99.98;	Pred. No. 0;		
Matches 1836; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

OY	8	GCAAGTCGGAGCTCGGAGCTGCTGCTTCTGTGTTCTCTTGCCACCCTGCCTGTGCCGGC	67
Db	1	GCAGGTCGGAGCTCGGAGCTGCTGCTTCTGTGTTCTCTTGCCACCCTGCCTGTGCCGGC	60
OY	68	TGCTTGGGCTGCCGAACAGACAGCGCTGGGCCACAGCACCTCAGAAGCCGACGAGCT	127
Db	61.	TGCTTGGGCTGCCGAACAGACAGCGCTGGGCCACAGCACCTCAGAAGCCGACGAGCT	120
OY	128	CGACGAGGGGCGCGCAGAGAGGTGGCGATCGCGTGTGAGGGCGCGCGGGCAGG	187
Db	121	CGACGAGGGGCGCGCAGAGAGGTGGCGATCGCGTGTGAGGGCGCGCGGGCAGG	180
OY	188	CGGGCGGGCGCCAGAGGGGGAAGAGCGGGGGCGGGGTCAGCCCCCTGGCCGGGGCGG	247
Db	181	CGGGCGGGCGCCAGAGGGGGAAGAGCGGGGGCGGGGTCAGCCCCCTGGCCGGGGCGG	240
OY	248	CGGGGGAATGTCGATGCTCTGACCGCATGCCGCTGCCCGGGGTGGGGGAGGAGCTGAAGCA	307
Db	241	CGGGGGAATGTCGATGCTCTGACCGCATGCCGCTGCCCGGGGTGGGGGAGGAGCTGAAGCA	300
OY	308	GGCCAAAGGAGATCGAGGAGCGCGAGAACTACTCTTCATGGCCACCCTCACCAGGGCGCC	367
Db	301	GGCCAAAGGAGATCGAGGAGCGCGAGAACTACTCTTCATGGCCACCCTCACCAGGGCGCC	360
OY	368	CAAGAACCAAATCCAGTTTGCATGATGACATGACAGAGTTCACCAAAATCCCCACCAAAAC	427
Db	361	CAAGAACCAAATCCAGTTTGCATGATGACATGACAGAGTTCACCAAAATCCCCACCAAAAC	420
OY	428	TGGCCGAAGATCTTTGTCTGCGCTCGATCTCACAGTCCCTCCACTGACAGCTACAGTTCAGC	487
Db	421	TGGCCGAAGATCTTTGTCTGCGCTCGATCTCACAGTCCCTCCACTGACAGCTACAGTTCAGC	480
OY	488	TGCATCTTACACAGATAGCTGTGATGATGAGGTTTCTCCCCGAGAGAACGACCAAAACCA	547
Db	481	TGCATCTTACACAGATAGCTGTGATGATGAGGTTTCTCCCCGAGAGAACGACCAAAACCA	540
OY	548	CTCCGAAGGCGCAGACCAATTCTGTGTGAAGAACATCAAGCAGCGAATTTGGACGCCG	607
Db	541	CTCCGAAGGCGCAGACCAATTCTGTGTGAAGAACATCAAGCAGCGAATTTGGACGCCG	600
OY	608	GGAGATTGAGATTGCAGAGACAAGACATGTCTGCTGTGATTTCACTCAGGAACGTGCTCA	667
Db	601	GGAGATTGAGATTGCAGAGACAAGACATGTCTGCTGTGATTTCACTCAGGAACGTGCTCA	660
OY	668	GGGGGAGAGCCCTTGGCTGTGCTTAATAATAGTGGGCTGTACACACATCACAGCCCAGAC	727
Db	661	GGGGGAGAGCCCTTGGCTGTGCTTAATAATAGTGGGCTGTACACACATCACAGCCCAGAC	720

QY	728	AGCGGTGTGATTGAGACACTCTGTGCCCTGGGGCTCAGTGCCGCTGCTCTGTGTA	787
Db	721	AGCGGTGTGATTGAGACACTCTGTGCCCTGGGGCTCAGTGCCGCTGCTCTGTGTA	780
QY	788	CATCTACTCACTCAGAAATGAAGTAGCTGCAGCACTGGCTGAGGCTGGAGTTGCAGTGT	847
Db	781	CATCTACTCACTCAGAAATGAAGTAGCTGCAGCACTGGCTGAGGCTGGAGTTGCAGTGT	840
QY	848	CGCTTGAAGGGCCGAGTCAGAAGATGACTTCTGTGTATTGACCCGCTGTGTGACAT	907
Db	841	CGCTTGAAGGGCCGAGTCAGAAGATGACTTCTGTGTATTGACCCGCTGTGTGACAT	900
QY	908	GGATGGGTGGCAGGCCAACATGATCCTGGATGATGGGGAGAGACTTAACCCACTGGCTTA	967
Db	901	GGATGGGTGGCAGGCCAACATGATCCTGGATGATGGGGAGAGACTTAACCCACTGGCTTA	960
QY	968	TAAGAAGTATCCAAACGTGTTTAAGAAGATCCGAGGCATTGTGGAAGAGAGCGTGA	1027
Db	961	TAAGAAGTATCCAAACGTGTTTAAGAAGATCCGAGGCATTGTGGAAGAGAGCGTGA	1020
QY	1028	TGTTACAGGCTGTATCAGCTCTCCAAACGTCGGGAAGCTCTGTGTCCGGCCATGAAG	1087
Db	1021	TGTTACAGGCTGTATCAGCTCTCCAAACGTCGGGAAGCTCTGTGTCCGGCCATGAAG	1080
QY	1088	CAATGATTTCTGTTACCAACACAGAAAGTTGATTAATCTGTACTGCTGCCGGAATCCATTT	1147
Db	1081	CAATGATTTCTGTTACCAACACAGAAAGTTGATTAATCTGTACTGCTGCCGGAATCCATTT	1140
QY	1148	GGATGGCTGAAGAGGACCACAGATGTGATGTTGGTGGGAACAAGTGTGTGTGG	1207
Db	1141	GGATGGCTGAAGAGGACCACAGATGTGATGTTGGTGGGAACAAGTGTGTGTGG	1200
QY	1208	CTATGTTAGGTAGGCAAGGGCTGCTGTGCTCTCAAAAGCTCTTGAGCAATGTGTA	1267
Db	1201	CTATGTTAGGTAGGCAAGGGCTGCTGTGCTCTCAAAAGCTCTTGAGCAATGTGTA	1260
QY	1268	CATTACCGAAATCGACCCCATCTGTGCTCTGACAGCCCTGCATGATGGGTTACAGGTGT	1327
Db	1261	CATTACCGAAATCGACCCCATCTGTGCTCTGACAGCCCTGCATGATGGGTTACAGGTGT	1320
QY	1328	AAAGCTAAATGAAGTCATCCGGCAAGTCGATGCTGTAATTAATCTGCACAGGAATAAGAA	1387
Db	1321	AAAGCTAAATGAAGTCATCCGGCAAGTCGATGCTGTAATTAATCTGCACAGGAATAAGAA	1380
QY	1388	TGTAGTGACACGGGAGCAGCTTGATCGCATGAAACAGTTGTATCGTATGCAATATGGG	1447
Db	1381	TGTAGTGACACGGGAGCAGCTTGATCGCATGAAACAGTTGTATCGTATGCAATATGGG	1440
QY	1448	CCACTCCAACACAGAAATCGATGTGACCAAGCCTCCGCACCTCCGAGCTGACGTGGAGCG	1507
Db	1441	CCACTCCAACACAGAAATCGATGTGACCAAGCCTCCGCACCTCCGAGCTGACGTGGAGCG	1500
QY	1508	AGTACGTTCTCAGGTGGAGCATGTCACTTGCCAGATGGCAACAGAGTTGCTCTCTGGC	1567
Db	1501	AGTACGTTCTCAGGTGGAGCATGTCACTTGCCAGATGGCAACAGAGTTGCTCTCTGGC	1560
QY	1568	AGAGGTGCTCTACTCAATTTGAGCTGCTCACAAGTTCCCACTTTGTTCTGTCCATCAC	1627
Db	1561	AGAGGTGCTCTACTCAATTTGAGCTGCTCACAAGTTCCCACTTTGTTCTGTCCATCAC	1620
QY	1628	AGCCACAACACAGGCTTTGGCACTGATAGACTCTATAATGCACCCGAGGGCCGATACAA	1687
Db	1621	AGCCACAACACAGGCTTTGGCACTGATAGACTCTATAATGCACCCGAGGGCCGATACAA	1680
QY	1688	GCAGAGTGTACTTGTCTCTTAAGAAAATGATGAATACGTTGCCAGCTTGCACTGCC	1747
Db	1681	GCAGAGTGTACTTGTCTCTTAAGAAAATGATGAATACGTTGCCAGCTTGCACTGCC	1740
QY	1748	ATCATTTGATGCCCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAA	1807
Db	1741	ATCATTTGATGCCCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAA	1800
QY	1808	CAAAAATGGGCCATTCAAAACCTAATTAATACAGATAC	1844

QY 1058 TGGGAAGCTCTGTGTCCGCCCATGACGTCATGATTTCTGTTACCAACACAGATTGTA 1117
|||||
Db 1021 TGGGAAGCTCTGTGTCCGCCCATGACGTCATGATTTCTGTTACCAACACAGATTGTA 1080
QY 1118 TAACCTGTACTGCTGCCGAGAAATCCATTTTGGATGGCCTGAAGAGGACACAGATGTGAT 1177
|||||
Db 1081 TAACCTGTACTGCTGCCGAGAAATCCATTTTGGATGGCCTGAAGAGGACACAGATGTGAT 1140
QY 1178 GTTTGGTGGAAACAAGTGTGTGTGTGTGGCTATGCTGAGGTAGGCAAGGGCTGTGTGC 1237
|||||
Db 1141 GTTTGGTGGAAACAAGTGTGTGTGTGTGTGGCTATGCTGAGGTAGGCAAGGGCTGTGTGC 1200
QY 1238 TGCTCTCAAGCTCTTGGAGCAATGTCTACATTAACGAATCGACCCCATCTGTCTCT 1297
|||||
Db 1201 TGCTCTCAAGCTCTTGGAGCAATGTCTACATTAACGAATCGACCCCATCTGTCTCT 1260
QY 1298 GCAGGCTGCATGATGGGTTCAGGGTGTAAAGCTAAATGAAGTCATCCGGCAAGTGA 1357
|||||
Db 1261 GCAGGCTGCATGATGGGTTCAGGGTGTAAAGCTAAATGAAGTCATCCGGCAAGTGA 1320
QY 1358 TGTCTAATACTTGCACAGGAATAAGATGATGACACGGGACACTTGTGATCGCAT 1417
|||||
Db 1321 TGTCTAATACTTGCACAGGAATAAGATGATGACACGGGACACTTGTGATCGCAT 1380
QY 1418 GAAAAACAGTGTGATGCTATGCAATATGGGCCACTCCAAACACAGAAATCGATGTGAC 1477
|||||
Db 1381 GAAAAACAGTGTGATGCTATGCAATATGGGCCACTCCAAACACAGAAATCGATGTGAC 1440
QY 1478 CCTCCGACCTCCGAGCTGACGTGGAGGAGTACGTTCTCAGGTGACCATGTCTCTG 1537
|||||
Db 1441 CCTCCGACCTCCGAGCTGACGTGGAGGAGTACGTTCTCAGGTGACCATGTCTCTG 1500
QY 1538 GCCAGATGGCAACGAGTGTCTCTCTGCGAGAGGGTCTCTACTCAATTGAGCTGCTC 1597
|||||
Db 1501 GCCAGATGGCAACGAGTGTCTCTCTGCGAGAGGGTCTCTACTCAATTGAGCTGCTC 1560
QY 1598 CACAGTCCCACTTTTGTCTGTCCATCACAGCCACACACAGAGGCTTTGGCACTGATAGA 1657
|||||
Db 1561 CACAGTCCCACTTTTGTCTGTCCATCACAGCCACACACAGAGGCTTTGGCACTGATAGA 1620
QY 1658 ACTCTAATAAGCAACCCGAGGGGCGATACACAGAGATGTGTACTGCTTCTTAAGAAAT 1717
|||||
Db 1621 ACTCTAATAAGCAACCCGAGGGGCGATACACAGAGATGTGTACTGCTTCTTAAGAAAT 1680
QY 1718 GGATGAATACGTTGCCAGCTTGCATCTGCCATCATTTGATGCCCACTTACAGAGCTGAC 1777
|||||
Db 1681 GGATGAATACGTTGCCAGCTTGCATCTGCCATCATTTGATGCCCACTTACAGAGCTGAC 1740
QY 1778 AGATGACCAACAATAATATCTGGGACTCAACAAATGGCCATTCAAACCTAATTATTA 1837
|||||
Db 1741 AGATGACCAACAATAATATCTGGGACTCAACAAATGGCCATTCAAACCTAATTATTA 1800
QY 1838 CAGATAC 1844
|||||
Db 1801 CAGATAC 1807

RESULT 6
HSU82761 2258 bp mRNA linear PRI 17-DEC-2001
LOCUS
DEFINITION Homo sapiens S-adenosyl homocysteine hydrolase homolog (XPVKona)
ACCESSION U82761
VERSION U82761.1 GI:2852124
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2258)
AUTHORS Cleaver,J.E., Afzal,V., Feeney,L., McDowell,M., Sadinski,W.,
Volpe,J.P.G., Busch,D.B., Coleman,D.M., Ziffer,D.W., Yu,Y.,
Nagaseawa,H. and Little,J.B.

TITLE Increased ultraviolet sensitivity and chromosomal instability
related to P53 function in the xeroderma pigmentosum variant
JOURNAL Cancer Res. 59 (5), 1102-1108 (1999)
MEDLINE 99168517
PUBMED 10070969
REFERENCE 2 (bases 1 to 2258)
AUTHORS Volpe,J.P.G., McDowell,M., Jostes,R.F., Afzal,V., Sadinski,W.,
Trask,B.J., Legeurs,R. and Cleaver,J.E.
TITLE Complementation of chromosomal instability in the xeroderma
pigmentosum variant by a gene on human chromosome 1 with homology
to S-adenosyl homocysteine hydrolase
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 2258)
AUTHORS Volpe,J.P.G., McDowell,M. and Cleaver,J.E.
TITLE Direct Submission
JOURNAL Submitted (19-DEC-1996) Dermatology, UCSF, 3rd and Parnassus, Box
0750, San Francisco, CA 94143, USA
FEATURES
source 1..2258
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/cell_type="fibroblasts"
/tissue_type="skin"
1..2258
/gene="XPVKona"
48..1550
/gene="XPVKona"
/note="Xeroderma pigmentosum variant"
/codon_start=1
/product="S-adenosyl homocysteine hydrolase homolog"
/protein_id="AAC01960.1"
/db_xref="GI:2852125"

BASE COUNT 603 a 513 c 542 g 600 t
ORIGIN
Query Match 83.9%; Score 1547; DB 9; Length 2258;
Best Local Similarity 100.0%; Pred. No. 1.8e-300;
Matches 1547; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 298 AGCTGAAGCAGGCCAAGAGATCGAGGACGCCGAGAAGTACTCTCTTCATGGCCACCGTCA 357
|||||
Db 1 AGCTGAAGCAGGCCAAGAGATCGAGGACGCCGAGAAGTACTCTCTTCATGGCCACCGTCA 60
QY 358 CCAAGGCGCCCAAGAACCAATCCAGTTTGTGATGACATGACAGAGTTCACCAAAATTCC 417
|||||
Db 61 CCAAGGCGCCCAAGAACCAATCCAGTTTGTGATGACATGACAGAGTTCACCAAAATTCC 120
QY 418 CCACCAAACTGCGCCGAAGATCTTGTCTGCTGATCTCACAGTCTCTCCACTGACAGCT 477
|||||
Db 121 CCACCAAACTGCGCCGAAGATCTTGTCTGCTGATCTCACAGTCTCTCCACTGACAGCT 180
QY 478 ACAGTTCAGCTGCATCTACACAGATAGCTGTGATGATGAGGTTTCTCCCGAGAAGAGC 537
|||||
Db 181 ACAGTTCAGCTGCATCTACACAGATAGCTGTGATGATGAGGTTTCTCCCGAGAAGAGC 240
QY 538 AGCAAACTCACTCCAAAGGCGACGACCAATTTCTGTGAAGAACATCAAGCAGGAGAT 597
|||||
Db 241 AGCAAACTCACTCCAAAGGCGACGACCAATTTCTGTGAAGAACATCAAGCAGGAGAT 300
QY 598 TTGAGCGCCGGAGATTGAGATTGACAGCAGACATGTCTGCTGTGATTTCACCTCAGGA 657
|||||
Db 301 TTGAGCGCCGGAGATTGAGATTGACAGCAGACATGTCTGCTGTGATTTCACCTCAGGA 360

OY	658	AACGTGCTCAGGGGGAGAAGCCCTTGGCTGGTGCTAAATAGTGGGCTGTACACATCA	717
Db	361	AACGTGCTCAGGGGGAGAAGCCCTTGGCTGGTGCTAAATAGTGGGCTGTACACATCA	420
OY	718	CAGCCAGACAGCGGTGTGATTGAGACACTGTGCCCTGGGGCTCAGTGCCTGGT	777
Db	421	CAGCCAGACAGCGGTGTGATTGAGACACTGTGCCCTGGGGCTCAGTGCCTGGT	480
OY	778	CTGCTTGTACATCTACTCACTCAGAATGAAGTAGCTGCAGCACCTGGCTGAGGCTGGAG	837
Db	481	CTGCTTGTACATCTACTCACTCAGAATGAAGTAGCTGCAGCACCTGGCTGAGGCTGGAG	540
OY	838	TTCAGATGTTCCGCTTGGAAAGGGCGAGTCAGAAGATGACTTCTGGTGTATTAACCGCT	897
Db	541	TTCAGATGTTCCGCTTGGAAAGGGCGAGTCAGAAGATGACTTCTGGTGTATTAACCGCT	600
OY	898	GTTGTAACATGGATGGGTGGCAGGCCAACATGATCTGGATGATGGGGAGACTTAACC	957
Db	601	GTTGTAACATGGATGGGTGGCAGGCCAACATGATCTGGATGATGGGGAGACTTAACC	660
OY	958	ACTGGGTTTATAAAGATATCCAACGTTGTTAAGAAGATCCGAGCATGTGGAAGAGA	1017
Db	661	ACTGGGTTTATAAAGATATCCAACGTTGTTAAGAAGATCCGAGCATGTGGAAGAGA	720
OY	1018	GCGTGACTGGTGTCAAGGCTGTATCAGCTCTCCAAGCTGGGAAGCTCTGTGTTCCGG	1077
Db	721	GCGTGACTGGTGTCAAGGCTGTATCAGCTCTCCAAGCTGGGAAGCTCTGTGTTCCGG	780
OY	1078	CCATGAACGTCATGATGATCTGTTACCAAACAGAAAGTTGATACTGTACTGCTGCCAG	1137
Db	781	CCATGAACGTCATGATGATCTGTTACCAAACAGAAAGTTGATACTGTACTGCTGCCAG	840
OY	1138	AATCCATTTTGGATGGCCTGAAGAGGACCACAGATGTGATGTTGGTGGAAACAAGTGG	1197
Db	841	AATCCATTTTGGATGGCCTGAAGAGGACCACAGATGTGATGTTGGTGGAAACAAGTGG	900
OY	1198	TGGTGTGTGGCTATGTTGAGGTAGGCAAGGGCTGCTGTGCTCTCAAAAGCTCTTGGAG	1257
Db	901	TGGTGTGTGGCTATGTTGAGGTAGGCAAGGGCTGCTGTGCTCTCAAAAGCTCTTGGAG	960
OY	1258	CAATGTCTACATTAACGAAATCGACCCCATCTGTGCTCTGCAGGCCCTGCATGATGGGT	1317
Db	961	CAATGTCTACATTAACGAAATCGACCCCATCTGTGCTCTGCAGGCCCTGCATGATGGGT	1020
OY	1318	TCAGGGTGTAAAGCTAAATGAAGTCATCCGGCAAGTCGATGTCGTAATAACTTGCACAG	1377
Db	1021	TCAGGGTGTAAAGCTAAATGAAGTCATCCGGCAAGTCGATGTCGTAATAACTTGCACAG	1080
OY	1378	GAAATAGAATGTAGTGACACAGGGAGCAGCTTGATCCGATGAAAAACAGTGTATCCGTAT	1437
Db	1081	GAAATAGAATGTAGTGACACAGGGAGCAGCTTGATCCGATGAAAAACAGTGTATCCGTAT	1140
OY	1438	GCAATATGGGCCACTTCCAACACAGAAATCGATGTGACCAAGCCTCCGCACCTCCGGAGCTGA	1497
Db	1141	GCAATATGGGCCACTTCCAACACAGAAATCGATGTGACCAAGCCTCCGCACCTCCGGAGCTGA	1200
OY	1498	CGTGGAGCGAGTACGTTCTCAGGTGAGCCATGTCATCTGGCCAGATGGCAACGAGTGG	1557
Db	1201	CGTGGAGCGAGTACGTTCTCAGGTGAGCCATGTCATCTGGCCAGATGGCAACGAGTGG	1260
OY	1558	TCCCTCGTGGCAGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTTCGCCACCTTGTTC	1617
Db	1261	TCCCTCGTGGCAGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTTCGCCACCTTGTTC	1320
OY	1618	TGTCATACACAGCCACAACACAGGCTTTGGCACATGATAGAATCTATATATGACACCCGAGG	1677
Db	1321	TGTCATACACAGCCACAACACAGGCTTTGGCACATGATAGAATCTATATATGACACCCGAGG	1380
OY	1678	GGCGATACAAGCAGATGTGTACTTGTCTCCTAAGAAAATGGATGAATACGTTGCCAGCT	1737
Db	1381	GGCGATACAAGCAGATGTGTACTTGTCTCCTAAGAAAATGGATGAATACGTTGCCAGCT	1440
OY	1738	TGCATCTGCCATCATTTGATGCCCACTTACAGAGCTGACAGATGACCAAGCAAAATATC	1797

Db	1441	TGCATCTGCCATCATTTGATGGCCACCCTTACAGAGCTGACAGATGACCAAGCAAAATATC	1500
QY	1798	TGGGACTCAACAAAAATGGGCCATTCAACCTAATTATACAGATAC	1844
Db	1501	TGGGACTCAACAAAAATGGGCCATTCAACCTAATTATACAGATAC	1547

RESULT	7
BC018218	
LOCUS	BC018218
DEFINITION	Mus musculus, S-adenosylhomocysteine hydrolase-like 1, clone MGC:18748 IMAGE:4007102, mRNA, complete cds.
ACCESSION	BC018218
VERSION	BC018218.1
KEYWORDS	GMC.
SOURCE	house mouse.
ORGANISM	Mus musculus

REFERENCE AUTHORS TITLE JOURNAL	
1 (bases 1 to 1772) Strausberg, R. Direct Submission Submitted (03-DEC-2001)	National Institutes of Health, Mammalian

REMARK	COMMENT
NIH-MGC Project URL: http://mgc.ncl.nih.gov	Contact: MGC help desk

Email: cgabbs-remall.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>
Series: IRAK Plate: 23 Row: n Column: 6
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Genomescan gene prediction, Similarity but not identity to protein.

FEATURES	Location/Qualifiers
source	1. .1772

CDS

```

/organism="Mus musculus"
/db_xref="taxon:10090"
/map="CZECH II"
/clone="MGC:18748 IMAGE:4007102"
/tissue_type="Mammary tumor metastatized to lung.
MTV-LTR/Mutl model. Expression driven by an MMTV-LTR
enhancer."
/clone_11b="NCI_CGAP_Lu30"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
191. 1693
/codon_start=1
/product="S-adenosylhomocysteine hydrolase-like 1"
/protein_id="AAH18218.1"
/db_xref="GI:17390493"
/db_xref="locusID:10768"
/translation="MATVTKAPKKQIQFADDMQEFTEKPTKTGRSLRSISOSSTD
YSSASATDSDDEVS PREKOQNTSKSSNCCVKNIKQAEFGREIEAEQDMASALIS
LRKRAAGEKPLAGAKIVGCTHITTAQTAVLLETLCALGAQCRWSACNTYSTONEVAAL
AEAGVASEATGKESEDEDFWMCIDRCVNMDCMOANMILDDGDLTHVYKKYPNVFKKI
RGIVAESVTGVRHLYOLSKAGRLCPAMVNDSTVKQFDNLVCCRSILDLGLKRTD
VMFGKQVVNVCVGVEVGKCCCAALKALGAIVYITTEIDPICALQACNDGFRVVKLNEVI

```


sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFZp564A1523) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de
Further information about the clone and the sequencing project is available at <http://www.mips.biochem.mpg.de/proj/cDNA/>.

FEATURES

Source

1. 2510

/organism="Homo sapiens"

/db_xref="taxon:9606"

/map="1, 15.55 CR from DIS248"

/clone="DKFZp564A1523"

/issue_type="Brain"

/clone_1lb="564 (synonym: hfbz2). Vector pAMP1; host

XL-2blue; sites NotI + SalI"

/dev_stage="fetal"

1. 1796

/gene="DKFZp564A1523"

<1. 1796

/note="S-adenosyl homocysteine hydrolase homolog XPYKona

splice variant"

/codon_start=3

/product="hypothetical protein"

/protein_id="CAB43223.1"

/db_xref="GI:4884204"

/translation="GLTVELDSSVCYILLRHRGSSRRKICAEISOPALPLPOE
TOAPSGESENLRNROGRRAESSONLSLRRLVSAOGRSDMRWMIQIMPEIDSL
ISCFLOIADDDQEFTEKPTKGRRLSRSISGSDTSYSAASYDSSDEVSPE
KQOTNSKSSNEFCVNIKQAEGRWSACNIYSTONEVAALAEAGVAFAMGSEDEDFW
THITAOAVLITLICALGAQCRWSACNIYSTONEVAALAEAGVAFAMGSEDEDFW
WCIDRCVNMDSQNMILLDDGDLTHVYKKYKPYKIRGIVESVTGHRLYQLSK
AGKLCVPAMNVDVTKOKFDNLVCCRESILDGKRTDVMGKQVNVCGYEVGKG
CCALKALGAIYITEIDPICALQACMDGFRVYKLVNEVIRQVDVITCTGNKVVYRE
HIDRMKNSCIVCMNGHNTSIDVTSLRTPELTWERVRSQVDHVIWPDGKRVYLAEGR
LNLSCSTVPTFYVITATTOALALIELYNAPERYKQDVYLLPKKMEYVASLHLP
FDAHLTELTDQAQKYLGLNKNGPKPNRYRY"

polya_site 2490
BASE COUNT 664 a 556 c 618 g 672 t
ORIGIN

Query Match 79.7%; Score 1469.4; DB 9; Length 2510;
Best Local Similarity 99.9%; Pred. No. 6.8e-285;
Matches 1470; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 374 GCAATCCAGTTGCTGATGACATGACAGAGTTCACCAATTCGCCACCAAACTGGCG 433
Db 323 GCAATCCAGTTGCTGATGACATGACAGAGTTCACCAATTCGCCACCAAACTGGCG 382
QY 434 AAGATCTTGTCTCGCTGATGACATGACAGTTCCTCAGTACAGTACAGTTCAGTGCATC 493
Db 383 AAGATCTTGTCTCGCTGATGACATGACAGTTCCTCAGTACAGTACAGTTCAGTGCATC 442
QY 494 CTACACAGATAGCTCTGATGATGAGGTTCTCCCGAGAGAGCAAACTCCAA 553
Db 443 CTACACAGATAGCTCTGATGATGAGGTTCTCCCGAGAGAGCAAACTCCAA 502
QY 554 GGGCAGCAGCAATTTCTGTGAGAACATCAAGCAGAGCAATTTGACGCGCGGAGAT 613
Db 503 GGGCAGCAGCAATTTCTGTGAGAACATCAAGCAGAGCAATTTGACGCGCGGAGAT 562
QY 614 TGAGATTCAGAGCAGACATGTCGCTCTGATTTCACTCAGGAAACGTCTCAGGGGGA 673
Db 563 TGAGATTCAGAGCAGACATGTCGCTCTGATTTCACTCAGGAAACGTCTCAGGGGGA 622
QY 674 GAAGCCTTGGCTGGTAAATAGTGGGCTGTACACACATCAGAGCCAGAGAGCGGT 733
Db 623 GAAGCCTTGGCTGGTAAATAGTGGGCTGTACACACATCAGAGCCAGAGAGCGGT 682
QY 734 GTTATGAGACACTGTGCTGGGCTCAGTGCAGCTGCTGCTGTACATCTA 793
Db 683 GTTATGAGACACTGTGCTGGGCTCAGTGCAGCTGCTGCTGTGTACATCTA 742

QY 794 CTCAACTCAGATGAGTAGTCTGACAGACTGGCTGAGGCTGAGTTGCAGTGTTCGCTTG 853
Db 743 CTCAACTCAGATGAGTAGTCTGACAGACTGGCTGAGGCTGAGTTGCAGTGTTCGCTTG 802
QY 854 GAAGGCGAGTCAGAGATGACTTCTGCTGCTGATTTGACCCGCTGTGTAACATGATG 913
Db 803 GAAGGCGAGTCAGAGATGACTTCTGCTGCTGATTTGACCCGCTGTGTAACATGATG 862

QY 914 GTGGCAGGCCAATGATGATCTGATGATGGGGAGACTTAACCCACTGGGTTATAGAA 973
Db 863 GTGGCAGGCCAATGATGATCTGATGATGGGGAGACTTAACCCACTGGGTTATAGAA 922

QY 974 GTATCCAAAGTGTTTAAGAGATCCAGGCAATTTGGAAGAGAGCGTACTGTTCA 1033
Db 923 GTATCCAAAGTGTTTAAGAGATCCAGGCAATTTGGAAGAGAGCGTACTGTTCA 982

QY 1034 CAGGCTGTATCAGCTCTCCAAAGCTGGGAGCTCTGTGTTCCGGCCATGAACGTCAATGA 1093
Db 983 CAGGCTGTATCAGCTCTCCAAAGCTGGGAGCTCTGTGTTCCGGCCATGAACGTCAATGA 1042

QY 1094 TTCTGTACCAACAGAGATTGATTAATCTGCTGCTGCCGAGAAATCCATTTTGATGG 1153
Db 1043 TTCTGTACCAACAGAGATTGATTAATCTGCTGCTGCCGAGAAATCCATTTTGATGG 1102

QY 1154 CCTGAGAGAGCACAGATGTGATGTTGGTGGGAAACAAGTGGTGTGTTGGCTATGG 1213
Db 1103 CCTGAGAGAGCACAGATGTGATGTTGGTGGGAAACAAGTGGTGTGTTGGCTATGG 1162

QY 1214 TGAGGTAGGCAAGGCTGCTGCTGCTCTCAAAAGCTCTTGAGAGCAATTTGCTACATTAC 1273
Db 1163 TGAGGTAGGCAAGGCTGCTGCTGCTCTCAAAAGCTCTTGAGAGCAATTTGCTACATTAC 1222

QY 1274 CGAATCGACCCCATCTGCTCTGCAAGCCTGCATGATGGGTTCAAGGCTGTAAGCT 1333
Db 1223 CGAATCGACCCCATCTGCTCTGCAAGCCTGCATGATGGGTTCAAGGCTGTAAGCT 1282

QY 1334 AATGAGTCAATCCGGCAAGTGCATGCTGTAATTAATTCACAGAGAAATAGAATAGT 1393
Db 1283 AATGAGTCAATCCGGCAAGTGCATGCTGTAATTAATTCACAGAGAAATAGAATAGT 1342

QY 1394 GACACGGGAGCACTTGGATCGCATGAAACAGTGTATCGTATGACATATGGCCACTC 1453
Db 1343 GACACGGGAGCACTTGGATCGCATGAAACAGTGTATCGTATGACATATGGCCACTC 1402

QY 1454 CAACACAGAAATGATGTGACCAAGCTCCGCACTCCGAGCTGACGTTGGAGCGAGTACG 1513
Db 1403 CAACACAGAAATGATGTGACCAAGCTCCGCACTCCGAGCTGACGTTGGAGCGAGTACG 1462

QY 1514 TTCTCAGGTGACCATGTCTATCTGGCCAGATGGCAACGAGTTGCTCCTGGCAGAGGG 1573
Db 1463 TTCTCAGGTGACCATGTCTATCTGGCCAGATGGCAACGAGTTGCTCCTGGCAGAGGG 1522

QY 1574 TCGTCTACTCAATTTGAGCTGCTCCACAGTTCCACCTTGTCTGTCATCACAGCCAC 1633
Db 1523 TCGTCTACTCAATTTGAGCTGCTCCACAGTTCCACCTTGTCTGTCATCACAGCCAC 1582

QY 1634 AACACAGGCTTTGGCACTGATAGAACTTAATATGACCCGAGGGGGGATACAAAGCAGGA 1693
Db 1583 AACACAGGCTTTGGCACTGATAGAACTTAATATGACCCGAGGGGGGATACAAAGCAGGA 1642

QY 1694 TGTGTACTTGTCTTCTAAGAAATGATGATAGTTCAGGCTTGCATCTGCAATCAT 1753
Db 1643 TGTGTACTTGTCTTCTAAGAAATGATGATAGTTCAGGCTTGCATCTGCAATCAT 1702

QY 1754 TGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAACAAAA 1813
Db 1703 TGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAACAAAA 1762

QY 1814 TGGGCCATTAACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1844
Db 1763 TGGGCCATTAACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1793

RESULT 9	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	REMARK	COMMENT
BC008349		2030 bp mRNA linear PRI 12-JUL-2001	BC008349					Homo sapiens, similar to S-adenosylhomocysteine hydrolase-like 1, clone IMAGE:3536052, mRNA, partial cds.					
BC008349			BC008349					1 (bases 1 to 2030)					
BC008349.1	GI:14249935							Strausberg, R.					
								Direct Submission					
								Submitted (25-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA					
								NIH-MGC Project URL: http://mgc.ncl.nih.gov					
								Contact: MGC help desk					
								Email: cgabs-remail.nih.gov					
								Tissue Procurement: DCTD/DP					
								CDNA Library Preparation: Rubin Laboratory					
								CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)					
								DNA Sequencing by: Genome Sequence Centre,					
								BC Cancer Agency, Vancouver, BC, Canada					
								info@cgsc.bc.ca					
								Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,					
								Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,					
								Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo					
								Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven					
								Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline					
								Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott,					
								Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,					
								George Yang, Scott Zuyderduyn, Marco Marra.					
								Clone distribution: MGC clone distribution information can be found					
								through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov					
								Series: IRAL Plate: 23 Row: 5 Column: 18.					
								Location/Qualifiers					
								1. 2030					
								/organism="Homo sapiens"					
								/db_xref="taxon:9606"					
								/clone="IMAGE:3536052"					
								/tissue_type="Lung, small cell carcinoma"					
								/clone_lib="NIH-MGC_7"					
								/lab_host="DH10B-R"					
								/note="Vector: POTB7"					
								<1.1838					
								/codon_start=3					
								/product="Similar to S-adenosylhomocysteine hydrolase-like					
								1"					
								/protein_id="AAH08349.1"					
								/db_xref="GI:14249936"					
								/translation="MSVQVVSAAAAPKEVELKDLSPSEAESQGLSTAAGVGAAP					
								AGGDPPEAPAPAAERPPVPGSPGAPALSPAGKVPQASAMKRSDPHHQHRDGG					

OY	149	GGTGGGCGATCGCGTGTCGGAGGGGCCCGCGCGGGCAGAGCCGGGGCGGCACAGAGGGGA	208
Db	134	GCGGGGCGGTGGAGACCCTGAGGCTCCAGCTCCCGCGGGAGCGGGCCCCCGGTCCTCCGG	193
OY	209	AAGAGCGGGGGCGGGGTCAGCCCGCTGGCCGGGGCGGGGGAATGTCATGCTGA	268
Db	194	CCCGGGCTCGGGGCCCGCGCTGCTCTCAGCCCCCGCGGGGAAGTGCTCAGCGCTC	253
OY	269	CGCATGCCGCTGCCCCGGGGTCGGGGAGAGAGCTGAAGCAAGCCCAAGAGATCGAAGACC	328
Db	254	GGCCATGAAGCGGAGCGCACCATCAACAGCACACGCGGCACCGCGCGCGAGCC	313
OY	329	CGAAGAATACTCCTTCATGGCCACCGTCACCAAGCGCC-----CAAGAAGCAAT	379
Db	314	CTGTGTACGCCCGAC--GGCACCGTCAACGAGCGCGCGGCACAGTCAAGAAGCAGAT	370
OY	380	CCAATTGCTGATGACATGCAAGAGTTCAACCAATTCCTACCAAACTGGCCGAAGATC	439
Db	371	CCAGTTGTCTGACACAGAAGCAAGTAATCAACAACGTCCACCAAATTTGACGTGCTC	430
OY	440	TTTTGCTCGCTCGATCTCACAGTCTCCACTGACAGCTFACAGTTCAAGTGCATCTTAC	499
Db	431	TTTGTCTCGTTCATTTCTCAGTCATCTACTGACAGCTFACAGCTCAGCGCTTCATFAC	490
OY	500	AGATAGCTCTGATGATGAGGTTCTCCCCGAGAGAAGCAGCAAACCACTCCAAGGGCAG	559
Db	491	AGATAGCTCTGATGATGAGACATCGCCCAGGGACACAGCAGCAAAAGAACTTAAGGMA	550
OY	560	CAGCAATTTCTGTGTGAAGAACATCAAGCAGGAGAAATTTGACCGCCGGAGATTGAGAT	619
Db	551	CAGTAGCTTCTGTGTTAAGAACATCAACACAGGACAGAGTTTGAAGCAAGAAATTGAAT	610
OY	620	TGCAGAGCAAGACATGTCTGCTCTGATTTCACTCAGAGAAAGCTGCTCAGGGGAGAGCC	679
Db	611	TGCTGAACAAGAAATGCTGCAATGATGGCTTTGAGGAGAGAGCTCAAGGAGAAAAGCC	670
OY	680	CTTGCTGTGCTAAATAGTGGGCTGTACACACATCACAGCCACAGCAGCGTGTGAT	739
Db	671	TTTGCTGGAGCCAAATCGTGGGTTGCACACACATCACAGCTCAGACTGCTGTGCTTAT	730
OY	740	TGAGACACTCTGTGCCCTGGGGGCTCAGTGCCGCTGCTCTGTGAACATCTACTCAAC	799
Db	731	GGAACCTCTGGGTGCTCTGGGGGCCAGTGCCGATGGGCTGCTGCAACATCTATTCCAC	790
OY	800	TCAGATGAAGTAGCTGCAGCAGCTGGCTGAGGCTGGAGTTGCAAGTGTGCTTGAAGSG	859
Db	791	TCATGAAGTAGGCTGCTGCTTAGCAGAAAGTGATTTCTGTTTTCCTGGAAGGG	850
OY	860	CGAGTCAGAAGATGACTTCTGTGTGTATTGACCCTGTGTGAACATGATGGGTGCA	919
Db	851	AGAGTCAGAAGATGACTTTTGTGTGTATCGATAGATGTGCAATGTGAGGGCTGCA	910
OY	920	GGCCACATGATCCTGGATGATGGGGGAGACTTAACCCACATGGGTTTATAAGAAGTACC	979
Db	911	GCCAAACATGATCTTGATGATGAGGGGATCTTACCACATGATTTATAAAAGTATCC	970
OY	980	AAACGTGTTTAAGAAGATCCGAGCAATTTGGAAGAGAGCGTGACTGTGTTCCAGGCT	1039
Db	971	CAACATGTTTAAGAAAAATCAAGGGCATAGTAGAGAGAGTGTTACTGAGTTCACAGGCT	1030
OY	1040	GTATCAGCTCTCCAAGCTGGAGACTCTGTGTTCCGGCCATGAACGTCAATGATTTCTGT	1099
Db	1031	GTACCACTGTCCAAAGCTGGAGCTGTGTGTTCCAGCCATGAATGTCATGACTCAGT	1090
OY	1100	TACCAACAGAAGTTTGATACTGTACTGCTGCCGAGAAATCCATTTTGATGGCTGAA	1159
Db	1091	CACCAACAGAAATTTGACAACCTCTACTGTGCGGTGAATCAATTTCTGATGACTTAA	1150
OY	1160	GAGGACCAAGATGTGATGTTTGTGGGAACAACAGTGTGTGTGTGGCTATGTTAGGT	1219
Db	1151	AAGGACAACGACATGATGTTTGTGTGGAAGCAAGTGTGTGTGTGTGGCTATGAGAGGT	1210

Db 603 CAGTACTTCTGTGTTAAGAACATCAACAGGAGAGTTGGACGAGAGAAATTGAAT 662
QY 620 TGCAGAGCAGACATGCTGCTCTGATTTCACTCAGGAAACGCTCAGGGGAGAGCC 679
Db 663 TGCTGAACAGAAATGCTGATGATGCTTTGAGAGAGAGAGCTCAAGAGAAAGCC 722
QY 680 CTGGCTGGTCTAAATAGTGGCTGTACACACATCAGCCAGACAGCGGTGTGAT 739
Db 723 TTGGCTGGAGCCAAATCGTGGGTGACACACATCAGCTCAGACTGCTGTCTAT 782
QY 740 TGAGACACTGCTGCTGCTGGGGCTGAGTGGCCGCTGCTGTGTACATCTACTAAC 799
Db 783 GGAACCTGCTGCTGCTGGGGCTGAGTGGCCGCTGCTGCTGCAACATCTATCCAC 842
QY 800 TCAGATGAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 859
Db 843 TCTCAATGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 902
QY 860 CGAGTCAGAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 919
Db 903 AGAGTCAGAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 962
QY 920 GGCCACATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 979
Db 963 GCCAATGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022
QY 980 AAAGCTGTTAAGAGATCCGAGCATTGCGAAGAGAGCGTACTGCTGCTGCTGCTGCTGCTGCT 1039
Db 1023 CAACATGTTAAGAAATCAAGGCGATGAGAGAGAGAGTGTACTGAGTTCACAGGCT 1082
QY 1040 GATCAGCTCTCCAAAGCTGGAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1099
Db 1083 GTACCACTGCTCCAAAGCTGGAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1142
QY 1100 TACCAACAGAGTGTGATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1159
Db 1143 CACCAACAGAGTGTGATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1202
QY 1160 GAGGACACAGTGTGATGTTGGTGGGAAACAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1219
Db 1203 AAGGACACAGATGATGTTGGTGGGAAACAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1262
QY 1220 AGGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1279
Db 1263 GGGAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1322
QY 1280 CGACCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1339
Db 1323 TGACCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1382
QY 1340 AGTCATCCGCGAAGTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1399
Db 1383 GGTATCCGACAGTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1442
QY 1400 GGAGCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1459
Db 1443 AGAGCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1502
QY 1460 AGAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1519
Db 1503 AGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1562
QY 1520 GGTGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1579
Db 1563 AGTGAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1622
QY 1580 ACTCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1639
Db 1623 GCTGAACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1682
QY 1640 GGCTTGGCACTGATGAGATCTATATGACCCGAGGGCGATACAGCAGATGTGTA 1699
Db 1683 GGCTCTTGGCTGATGAGATCTTACATGCTGCTGAGGGTGGCTATAGCAGAGATGTCTA 1742

QY 1700 CTGCTCTCCAGAAATGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1759
Db 1743 CCTGTTGCCAGAGATGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1802
QY 1760 CCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAACAAAATGGGCC 1819
Db 1803 CCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAACAAAATGGGCC 1862
QY 1820 ATCAACCTAATTATACAGATA 1843
Db 1863 CTCAAGCTAATTACTACAGATA 1886

RESULT 11
AB020635
LOCUS
DEFINITION Homo sapiens mRNA for KIAA0828 protein, partial cds.
ACCESSION AB020635
VERSION AB020635.1 GI:4240144
KEYWORDS
SOURCE Homo sapiens adult male brain cDNA to mRNA, clone_11b:pluescriptII
SK plus clone:hh04230.

ORGANISM Homo sapiens
REFERENCE 1 (sites)
AUTHORS Nagase,T., Ishikawa,K., Suyama,M., Kikuno,R., Hirose,M., Miyajima,N., Tanaka,A., Kotani,H., Nomura,N. and Ohara,O.
TITLE Prediction of the coding sequences of unidentified human genes. XII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro
JOURNAL DNA Res. 5 (6), 355-364 (1998)
MEDLINE 99156230
REFERENCE 2 (bases 1 to 5025)
AUTHORS Ohara,O., Suyama,M., Kikuno,R., Nagase,T. and Ishikawa,K.
TITLE Direct Submission
JOURNAL Submitted (02-DEC-1998) Osamu Ohara, Kazusa DNA Research Institute, Laboratory of DNA Technology; Yana 1532-3, Kisarazu, Chiba 292-0812, Japan (E-mail:cdna@info.kazusa.or.jp, Tel: +81-438-52-3913, Fax: +81-438-52-3914)

FEATURES
source
1. 5025
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="hh04230"
/sex="male"
/tissue_type="brain"
/clone_11b="pluescriptII SK plus"
/dev_stage="adult"
1. 5025
/gene="KIAA0828"
<1. 1860
/gene="KIAA0828"
/codon_start=1
/product="KIAA0828 protein"
/protein_id="BAA74851.1"
/db_xref="GI:4240145"

translation="EPVVAEEVMSVQVNSAAAKVPEVELKDISPEAESQIGLST
AAGVAMPAGGDPAPAPAPAPVPPGSGPAALSPAGKVPQASAMKRSDPH
OHORHRDGEALVSPDGTVEAPRTVKIOFADQKOEFNKRPKIGRSLRSIOS
TDSYSSAASYTSSDDETSRDKQKNSKSSDFCVKNIKQAEGRREIEIAEQEMPA
LMAIRKRAOGEKPLAGAKIVGCTHITAQTAVLMETIGALGACRMAACNISTLNEVA
AAIAESGEFVFAWKGESEDEDFWICIDRCVNEGOMNILDGDLTHWYKKYPNM
KKIKGIVESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQFDNLKYCRESILDLKR
TIDMFEKGKOVVCGYGEVGGCCALKAMGSIVYVEIDPICALQACMDGFRIVKLN
EYIRQVDIVITCTGNKNVYTRHIDRMKNSCIVCNMGHSNTEIDVSLRTPELTWERV
RSQVDHYIMPDKRIYVLAEGRLNLSCSTVPTFVLSIRATQALALIELYNAPEGRY
KQDYVLLPKMDEYVASLHLPFDALHLELTDEQAKYLGKNGPFPKPNYYR"

BASE COUNT 1281 a 1179 c 1136 g 1429 t
ORIGIN
Query Match 54.0%; Score 995.6; DB 9; Length 5025;

COMMENT
NEO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing: Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).

FEATURES
Source Location/Qualifiers
1..1987
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="COLF0094"
/tissue_type="colon mucosa"
/clone_lib="COLF"
/note="cloning vector PME18SFL3"

BASE COUNT 557 a 430 c 501 g 499 t
ORIGIN

Query Match 53.1%; Score 979.8; DB 9; Length 1987;
Best Local Similarity 78.9%; Pred. No. 1.4e-186;
Matches 1167; Conservative 0; Mismatches 312; Indels 0; Gaps 0;

```
QY 365 GCCCAAGAGCAAAATCCAGTTTGTGATGACATGACAGAGTTTCCACCAATTCGCCACCAA 424
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 316 GCCTGAGTGATGATCCAGTTTGTGATGACATGACAGAGTTTCCACCAATTCGCCACCAA 375

QY 425 AACTGGCCGAGAAATCTTGTCTCGCTCGATCTCAGATCTCTCCACTGACAGTTACAGTTC 484
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 376 AATTGACGTCGCTCTTGTCTCGTCCATTTCTCAGTCTACTGACAGTTACAGTTC 435

QY 485 AGCTGCATCTTACACAGATAGCTCTGATGATGAGTTTCTCCCGAGAGAGCAAGCAAAAC 544
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 436 AGCGGCTTCATATACAGATAGCTCTGATGATGAGATCGCCAGGAGCAAGCAAAAC 495

QY 545 CAACTCCAGAGGAGCAGCAATTTCTGTGAGAAACATCAAGCAGGAGCAAGATTGGAGC 604
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 496 GAACCTTAAGGAGCAGTACTTCTGTGTAGAACATCAACAGGAGCAAGTTTGGAGC 555

QY 605 CCGGAGATGAGATTCAGAGCAAGACATGCTCTGATTTCACTCAGAAACGTGC 664
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 556 AAGAGAAATGAATGTGTAACAAGAAATGCTGATTTGATGGCTTTGAGAGAGAGAGC 615

QY 665 TCAGGGGAGAGCCCTTGCTGTGCTAAATAGTGGCTGTACACATCAGACCCCA 724
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 616 TCAGAGAGAAAGCCTTTGGCTGGAGCCAAATCGTGGTTCACACATCACTGCTCA 675

QY 725 GACAGCGGTGTGATGAGACACTCTGCTGCTGGGGCTCAGTGGCTGTGCTGCTTG 784
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 676 GACTGCTGCTTATGAAACTCTGGGCTCTGGGGCCAGTGGCATGGGCTGCCTG 735

QY 785 TAACATCTACTCAACTCAGATGAGTAGCTGACACTGCTGAGCTGAGAGTGCAGT 844
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 736 CAACATCTATTCACACTCTCAATGAGAGGCTGCTCTAGCAGAAAGTGATTTCTCTGT 795

QY 845 GTTGGCTTGAAGGCGAGTCAAGAGATGACTTCTGTGTATTTGACCGCTGTGTGAA 904
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 796 TTTTGCTTGAAGGAGAGAGTCAAGATGACTTCTGTGTGTATGATGAGAGGATCTTACC 855

QY 905 CATGATGGGTGGCAGGCCAATGATCTCGATGATGGGGGAGACTTAACCCACTGGGT 964
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 856 TGTGAGGGCTGGCAGGCCAATGATCTTGTGATGAGAGGATCTTACCCTGAT 915

QY 965 TTATAGAGATATCCAAACGTTTAAAGAGATCCGAGCATTTGGAAGAGAGCGTGAC 1024
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 916 TTATAGAGATATCCCAACATGTTTAAAGAAATCAAGGCGATAGTAGAGAGAGTGTAC 975

QY 1025 TGGTGTTCACAGGCTGTATCAGCTTCCAAAGCTGGAGAGCTGTGTTCGGCCATGAA 1084
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 976 TGAAGTTCACAGGCTGTATCAGCTTCCAAAGCTGGAGAGCTGTGTTCAGCCATGAA 1035

QY 1085 CGTCAATGATTTCTGTACCAACAGAGATTGATACTGTGCTGCCGAGATCCAT 1144
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 1036 TGTCAATGACTCAGTCACCAACAGAAATTTGACAACCTCTACTGTTGCCGTGAATCAT 1095
QY 1145 TTGGATGGCCTGAAGAGAGCACCACAGATGTGATGTTGGTGGGAACAAGTGTGTG 1204
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1096 TCTTGATGAGCTTAAAGAGACACAGACATGATGTTGTGGAAGCAAGTGTGTG 1155
QY 1205 TGGCTATGTTGAGTAGGCAAGGCTGCTGTGCTGCTCTCAAGCTCTTGAGCAATGT 1264
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1156 TGGCTATGAGAGAGTGGGGAAGGGTGTGCTGTGCTGCCCTGAAGCCATGGGCTCCAT 1215
QY 1265 CTACATTTACCGAATCGACCCCATCTGTGCTCTGACAGGCTTCATGATGGTTCAGGT 1324
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1216 GTATGTAAGTGAATGACCCCATCTGTGCTGCCCTGAAGCCATGATGATGATTTGACT 1275
QY 1325 GGTAAAGCTAAATGAAGTATCCGGCAAGTGCATGCTAATTAATTGACAGAAATTA 1384
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1276 GGTGAATTAATGAGGTCAATCCGCAAGTGCATGATGATGATGATGATGATGATGAT 1335
QY 1385 GAATGTAGTACACGGAGCACTTGGATGCTGATGAATAACAGTTGATGATGATGAT 1444
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1336 GAATGTGTAACAGAGAGCACTTGGACCGTATGAAGATAGCTGATGCTGTTGTACAT 1395
QY 1445 GGGCCACTCCAAACACAGAAATGATGTGACAGGCTCCGCACTCCGAGAGTACAGTGG 1504
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1396 GGGACATTTCCAAACACAGAAATGATGTGACGTTGGAGTCTGCGGACACCAAGACTGAC 1455
QY 1505 GCGAGTACGTTCTCAGTGTGACCATGTCATCTGCGCAGATGGCAACAGAGTTGCTGCT 1564
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1456 GCGAGTACATCTCAGTGTGACCATGTCATGATGATGATGATGATGATGATGATGAT 1515
QY 1565 GCGAGAGGCTGCTCAGTGTGACCATGTCATGATGATGATGATGATGATGATGATGAT 1624
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1516 GCGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1575
QY 1625 CACAGCCACACACAGGCTTTGGCAGTGTGATGATGATGATGATGATGATGATGAT 1684
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1576 CACTGCTACTACTCAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1635
QY 1685 CAAGCAGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1744
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1636 TAAGCAGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1695
QY 1745 GCCATCATTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1804
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1696 GCCTACCTTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1755
QY 1805 CAACAAAATGGGCCATTTCAACCTTAATTAACAGATA 1843
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1756 CAACAGAATGGGCCCTTCAAGCCTAATTAACAGATA 1794
```

RESULT 13
AY113501
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

AY113501 1792 bp mRNA linear INV 21-MAY-2002
Drosophila melanogaster RE58316 full insert cDNA.
AY113501
AY113501.1 GI:21064552
FLI_CDNA.
Drosophila melanogaster.
Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1792)
Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Carlson, J.,
Champe, M., Chavez, C., Dorsett, V., Dresnek, D., Farfan, D., Fritze, E.,
George, R., Gonzalez, M., Guarin, H., Kronmiller, B., Li, P., Liao, G.,
Miranda, A., Mungall, C. J., Nuno, J., Parag, V., Park, S.,
Patel, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Rubin, G. M.
and Ceiniker, S.
Direct Submission
Submitted (16-MAY-2002) Berkeley Drosophila Genome Project,
Lawrence Berkeley National Laboratory, One Cyclotron Road,

QY 1825 AACCTAATTAATACAGATA 1843
DB 1677 AAGCCAATTAATACAGATA 1695

RESULT 14
AF035319

LOCUS AF035319 1369 bp mRNA linear PRI 07-JUL-2000
DEFINITION Homo sapiens clone 23931 mRNA, partial cds.
ACCESSION AF035319
VERSION AF035319.1 GI:2661082
KEYWORDS FLI_CDNA.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS 1 (bases 1 to 1369)
Anderson, B., Wentland, M.A., Ricafrente, J.Y., Liu, W. and Gibbs, R.A.
TITLE A 'double adaptor' method for improved shotgun library construction
JOURNAL Anal. Biochem. 236 (1), 107-113 (1996)
MEDLINE 96207227
PubMed 8619474

REFERENCE
AUTHORS 2 (bases 1 to 1369)
Yu, W., Anderson, B., Worley, K.C., Muzny, D.M., Ding, Y., Liu, W.,
Ricafrente, J.Y., Wentland, M.A., Lennon, G. and Gibbs, R.A.
TITLE Large-scale concatenation cDNA sequencing
JOURNAL Genome Res. 7 (4), 353-358 (1997)
MEDLINE 97264341
PubMed 9110174

REFERENCE
AUTHORS 3 (bases 1 to 1369)
Yu, W., Sarginson, J. and Gibbs, R.A.
TITLE Direct Submission
JOURNAL Submitted (20-NOV-1997) Molecular and Human Genetics, Baylor
College of Medicine, One Baylor Plaza S930, Houston, TX 77030, USA

FEATURES
source location/Qualifiers
1. 1369
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="23931"
/sex="female"
/tissue_type="brain"
/clone_lib="INIB"
/dev_stage="infant"
/note="This clone is similar to Drosophila M. Abd-B gene
encoded by GenBank Accession Number L07835. The I.M.A.G.E.
Consortium clone ID number is 23931"
1. 657
/note="similar to S-adenosylhomocysteine hydrolase"
/codon_start=1
/protein_id="AAB88189.1"
/db_xref="GI:2661083"
/translation="PRAMCGYGEVKGCCCAALKALGAIVYITEIDICALQACMDGFR
VYKLEIVROVDVITCTGCKNVTREHLDKMKNSCIYCNMGHSNTEIDVTSRPEL
TWERVRSQVDHVIWPDGKRYVLLAEGRLNLSCSTVPFVLSITATTQALALIELYNA
PEGRYKQDVYLLPKMKDEYVASLHLPFDAHLELTDDQAKYVLGNKNGPFKNYRY

CDS
BASE COUNT 370 a 314 c 286 g 399 t
ORIGIN

Query Match 34.9%; Score 644; DB 9; Length 1369;
Best Local Similarity 100.0%; Pred. No. 3.9e-119;
Matches 644; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1201 TGTGTGGCTATGGTAGGAGGAGGCTGTGCTCTCAAGCTCTTGAGCAA 1260
DB 11 TGTGTGGCTATGGTAGGAGGAGGCTGTGCTCTCAAGCTCTTGAGCAA 70

QY 1261 TTGCTACATTAACGAATCGACCCCATCTGTCTCTGCAGGCTGCATGATGGTTCA 1320
DB 71 TTGCTACATTAACGAATCGACCCCATCTGTCTCTGCAGGCTGCATGATGGTTCA 130

QY 1321 GGTGTAAAGCTAAATGAAGTCATCCGGCAAGTCGATGCTAATACTTGACAGGAA 1380
DB 131 GGTGTAAAGCTAAATGAAGTCATCCGGCAAGTCGATGCTAATACTTGACAGGAA 190

QY 1381 ATAGAATGTAGTGACACGGAGACACTTGATCGCATGAAAAACAGTTGTATCGTATGCA 1440
DB 191 ATAGAATGTAGTGACACGGAGACACTTGATCGCATGAAAAACAGTTGTATCGTATGCA 250

QY 1441 ATATGGGCCCTCCACACAGAAATCGATGTGACCAAGCCTCCGCACTCCGAGCTGACGT 1500
DB 251 ATATGGGCCCTCCACACAGAAATCGATGTGACCAAGCCTCCGCACTCCGAGCTGACGT 310

QY 1501 GGGAGCGAGTACGTTCTCAGGTGGACCATGTCATCTGGCCAGATGGCAACGAGTTGCC 1560
DB 311 GGGAGCGAGTACGTTCTCAGGTGGACCATGTCATCTGGCCAGATGGCAACGAGTTGCC 370

QY 1561 TCCTGGCAGAGGCTGCTCTACTCAATTTGAGCTGCTCCACAGTTCCACCTTTGTTCTGT 1620
DB 371 TCCTGGCAGAGGCTGCTCTACTCAATTTGAGCTGCTCCACAGTTCCACCTTTGTTCTGT 430

QY 1621 CCATCACAGCCACAACACAGAGCTTTGGCAGTATAGAACTCTATAATGACCCGAGGGGC 1680
DB 431 CCATCACAGCCACAACACAGAGCTTTGGCAGTATAGAACTCTATAATGACCCGAGGGGC 490

QY 1681 GATACACAGCAGATGTGTACTTGTCTCTTACAAAATGGATGATATACGTTGCCAGCTTGC 1740
DB 491 GATACACAGCAGATGTGTACTTGTCTCTTACAAAATGGATGATATACGTTGCCAGCTTGC 550

QY 1741 ATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1800
DB 551 ATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 610

QY 1801 GACTCAACAAAATGGGCCATTCAACCTAATATTACAGATAC 1844
DB 611 GACTCAACAAAATGGGCCATTCAACCTAATATTACAGATAC 654

RESULT 15
BC003631 1323 bp mRNA linear PRI 12-JUL-2001
LOCUS BC003631
DEFINITION Homo sapiens, similar to S-adenosylhomocysteine hydrolase-like 1,
clone IMAGE:3010755, mRNA, partial cds.
ACCESSION BC003631
VERSION BC003631.1 GI:13177702
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1323)
Strausberg, R.
Direct Submission
Submitted (26-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
contact: amadansystemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahay, Stephanie Ford, Julia
Greene, Mark Kellerman and Anuradha Madan

REMARK
COMMENT
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 3 Row: B Column: 6.
location/Qualifiers
1. 1323
/organism="Homo sapiens"
/db_xref="taxon:9606"

/clone="IMAGE:3010755"
/tissue_type="Muscle, rhabdomyosarcoma"
/clone_1ib="NIH_MGC_17"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
<1..611
/codon_start=3
/product="Similar to S-adenosylhomocysteine hydrolase-like
1"

/protein_id="AAH03631.1"
/db_xref="GI:13177703"
/translation="HEEALGAIYVITEIDPICALQACMDGFRVYKLNVEYIRQVDVIT
CTGNKNVYTRHLDKMKNSCIVCMGHSNTEIDVTSLRPELTWERVRSQVDHVIWD
GKRVVLAEGRLNLSCSTVPTFVLSITATQALALIELYNAPGRIYKQDYVLLPKM
DEVVASLHLPFDALHLELTDQAKYLGILNKNKNGPFKNYRY"

BASE COUNT 365 a 303 c 271 g 384 t
ORIGIN

Query Match 32.5%; Score 599; DB 9; Length 1323;
Best Local Similarity 100.0%; Pred. No. 4.2e-110;
Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1246 AAGCTCTTGAGCAATGTCATACATTAACGAATCGACCCCATCTGTGCTCGACGGCCT 1305
DB 10 AAGCTCTTGAGCAATGTCATACATTAACGAATCGACCCCATCTGTGCTCGACGGCCT 69
QY 1306 GCATGGATGGGTCAGGGGTGTAAGCTAAATGAAGTCATCCGCAAGTCGATGTCCTAA 1365
DB 70 GCATGGATGGGTCAGGGGTGTAAGCTAAATGAAGTCATCCGCAAGTCGATGTCCTAA 129
QY 1366 TAACTTGACACAGAAATAAGATAGTAGACACAGGAGACACTTGATCGCATGAAAAACA 1425
DB 130 TAACTTGACACAGAAATAAGATAGTAGACACAGGAGACACTTGATCGCATGAAAAACA 189
QY 1426 GTGTATCGTATGCAATATGGGCCACTCCACACACAGAAATCGATGTGACCAAGCCTCCGCA 1485
DB 190 GTGTATCGTATGCAATATGGGCCACTCCACACACAGAAATCGATGTGACCAAGCCTCCGCA 249
QY 1486 CTCGGAGCTGACGTGGAGCCACTACGTTCTCAGGTGGACCATGTCATCTGGCCAGATG 1545
DB 250 CTCGGAGCTGACGTGGAGCCACTACGTTCTCAGGTGGACCATGTCATCTGGCCAGATG 309
QY 1546 GCAAAAGAGTGTCTCTCGGAGAGGGTCTCTACTCAATTGAGCTGCTCCACAGTTC 1605
DB 310 GCAAAAGAGTGTCTCTCGGAGAGGGTCTCTACTCAATTGAGCTGCTCCACAGTTC 369
QY 1606 CCACTTTGTTCTGTCCATCACAGCCACAACACAGGCTTTGGCACTGATAGAACTCTATA 1665
DB 370 CCACTTTGTTCTGTCCATCACAGCCACAACACAGGCTTTGGCACTGATAGAACTCTATA 429
QY 1666 ATGACCCGAGGGCGATACAGAGAGATGTACTTGTCTCTAAGAAATGATGAAT 1725
DB 430 ATGACCCGAGGGCGATACAGAGAGATGTACTTGTCTCTAAGAAATGATGAAT 489
QY 1726 ACGTTGCCAGCTTGCAATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACC 1785
DB 490 ACGTTGCCAGCTTGCAATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACC 549
QY 1786 AAGCAAAATATCTGGGACTCAACAAAATGGGCATTCAAACCTAATTATTACAGATAC 1844
DB 550 AAGCAAAATATCTGGGACTCAACAAAATGGGCATTCAAACCTAATTATTACAGATAC 608

Search completed: April 21, 2003, 02:39:39
Job time : 4773.08 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2003, 23:31:25 ; Search time 2489.43 Seconds
(without alignments)
11996.532 Million cell updates/sec

Title: US-09-782-051-1_COPY_1_1844

Perfect score: 1844

Sequence: 1 ggcgcggcgcagtcggagct.....aacctaattattacagatac 1844

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_hnv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	963.4	52.2	3234	11 AK014539	AK014539 Mus muscu
2	926.2	50.2	938	9 AL527928	AL527928 AL527928
3	897.4	48.7	921	9 AL551097	AL551097 AL551097
4	873	47.3	891	9 AL520704	AL520704 AL520704
5	862.6	46.8	1042	13 BM449470	BM449470 AGENCOURT
6	856.2	46.4	865	9 AL519571	AL519571 AL519571

7	847	45.9	967	13 BM468564	BM468564 AGENCOURT
8	845.8	45.9	858	9 AL554850	AL554850 AL554850
9	818.2	44.4	1042	13 BM452485	BM452485 AGENCOURT
10	814	44.1	839	9 AL525407	AL525407 AL525407
11	807.4	43.8	867	13 BM453920	BM453920 AGENCOURT
12	775.6	42.1	871	12 BG708016	BG708016 602671341
13	773.4	41.9	928	14 BQ641487	BQ641487 AGENCOURT
14	762.2	41.3	857	13 B1668412	B1668412 603292688
15	732.8	39.7	764	9 AU126038	AU126038 AU126038
16	732.2	39.7	894	14 BQ896777	BQ896777 AGENCOURT
17	730.4	39.6	748	9 AU124992	AU124992 AU124992
18	724.4	39.3	927	14 BQ894873	BQ894873 AGENCOURT
19	723.4	39.2	843	13 B1457815	B1457815 603198477
20	720.4	39.1	933	14 BQ889173	BQ889173 AGENCOURT
21	716.4	38.9	747	12 BG697210	BG697210 602660444
22	716.4	38.9	832	10 BE270773	BE270773 600943931
23	713.2	38.7	744	12 BG699160	BG699160 602678818
24	712.2	38.6	806	12 BG774719	BG774719 602662763
25	706.2	38.3	741	13 B1333908	B1333908 602999453
26	703.8	38.2	969	12 BG259448	BG259448 602378544
27	703	38.1	1071	13 BM548493	BM548493 AGENCOURT
28	702.4	38.1	801	13 B1335041	B1335041 602999053
29	698.2	37.9	751	9 AU133711	AU133711 AU133711
30	697.2	37.8	801	13 B1551474	B1551474 603194313
31	695.2	37.7	785	12 BG699237	BG699237 602678908
32	692.4	37.5	745	9 AU131962	AU131962 AU131962
33	692	37.5	773	13 B1549142	B1549142 603189237
34	689.4	37.4	1088	12 BF797186	BF797186 602257870
35	687.8	37.3	691	14 BM837838	BM837838 K-EST0113
36	681.6	37.0	767	12 BG699021	BG699021 602678652
37	673	36.5	1098	13 BM454352	BM454352 AGENCOURT
38	672.6	36.5	1174	14 BM927200	BM927200 AGENCOURT
39	672.4	36.5	700	9 AU133884	AU133884 AU133884
40	672.4	36.5	956	12 BF792750	BF792750 602253481
41	669.8	36.3	749	12 BG700812	BG700812 602681707
42	667.6	36.2	713	13 B1462766	B1462766 603205724
43	665.4	36.1	802	12 BG704022	BG704022 602687151
44	661.8	35.9	930	12 BG751929	BG751929 602731492
45	661	35.8	769	13 B1551275	B1551275 603194678

ALIGNMENTS

RESULT 1 AK014539

LOCUS AK014539 3234 bp mRNA linear HTC 19-JAN-2002

DEFINITION Mus musculus 0 day neonate skin cDNA, RIKEN full-length enriched library, clone:4631427C17:homolog to PUTATIVE ADENOSYLHOMOCYSTEINASE (EC 3.3.1.1) (S-ADENOSYL-L-HOMOCYSTEINE HYDROLASE) (ADHCYASE), full insert sequence.

ACCESSION AK014539
VERSION AK014539.1 GI:12852459
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) 0 day neonate skin cDNA to mRNA, clone_11b:RIKEN full-length enriched mouse cDNA library
clone:4631427C17.

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE

1 Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

PUBMED 99279253

MEDLINE 10349636

AUTHORS

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itch, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL

MEDLINE 20499374

OY	798	ACTCAGAAATGAAGTACCTGTCAGCACTGGCGCTGAGGCCTGGAGTTGCAGTGTTCCGCTTGGAA	857
Db	836	ACCCTCAATGAAGTGGCTGCTGCTCTCGCAGAAGTGGGTTCGCCGCTTTGCCCTGGAAA	895
OY	858	GCGAGTCAAGAGATGACTTCTGGTGTATGTAACCGCTGTGTGAACAAGATGGGTGG	917
Db	896	GGAGATCAGAGATCACTTTTGGTGTGATGACAGATGCGTGAATGTGGAGGGCTGG	955
OY	918	CAGGCCAACATGATTCCTGGATGATGGGGAGACTTAAACCACCTGGGTTATAAAGATAT	977
Db	956	CAGCCAAACATGATTAAGTATGATGATGAGGAGATCTGACTCACCTGGATTATTAAGATAT	1015
OY	978	CCAAACGTGTTTAAAGATATCCGAGGCATTTGGAAAGAGAGCGTGAAGTGTTCACAGG	1037
Db	1016	CCCACATGTTTAAAAAATCAAAAGCATAGTCGAGAGAGTGTACTGGAGTCCATAGG	1075
OY	1038	CTGTATCAGCTCTCCCAAAGCTGGGAAGCTCTGTGTTCCGGCCATGAACGTCAATGATCT	1097
Db	1076	CTGTACCAACTGTCCAAGCTGGGAAGCTGTGTGTTCCCGCATGAACGTCAATGACTCA	1135
OY	1098	GTTACCAAAACAGAGTTGATACTTGTACTGCTGCCGAGAAATCCATTTTGGATGGCCTG	1157
Db	1136	GTCACATAAGCAGAAATTTGATAACCTCTACTGTGGCGTGAATCTATCTTGACGCACTT	1195
OY	1158	AAGAGGACCACAGATGTCATGTTTGGTGGGAACAAGTGGTGTGTGGCTATGGTGAG	1217
Db	1196	AAAAGGACACAGATATGATGTTTGGTGGAAAGCAGGTGGTGTGTGGCTATGGAGAG	1255
OY	1218	GTAGGCAAGGGCTGCTGTGCTGCTCTCTCAAAGCTCTTGSAGCAATGTCTACATTAACGAA	1277
Db	1256	GTGGAAAAGGGGTGCTGGCTGCTCTGAAAGCCATGGGCTCCATTGTGTACGTGACTGAG	1315
OY	1278	ATCGACCCCATCTGTGCTCTGCAGGCGCTGCATGATGGGTTACAGGGTGAAGCTAAAT	1337
Db	1316	ATTGACCCCATCTGTGCTCTGCAGGCGCTGATGATGATGATTCGACTAGTGAAGCTGAAT	1375
OY	1338	GAAATCATCCGGCAAGTGCATGTCGTAATAACTGACAGGAATAAGATGTAGTACA	1397
Db	1376	GAAATCATCCGACAGTGCAGACATGTTTATTAATCTGCACAGGGAACAAGATGTGTAACC	1435
OY	1398	CGGAGCAGCTTGGATCGCATGA AAAACAGTTGTATCGTATGCAATATGGGCCACTCCAAC	1457
Db	1436	AGAGAGCAGCTTGGACCCGATGAAGAATAGCTGCATCGTTTGTAACTGGGCCATTC AAC	1495
OY	1458	ACAGAAATCGATGTGACAGCCCTCCGCACTCCGGAGCTGACGTGGAGCGAGTACG-TTC	1516
Db	1496	ACCGAAATGTATGTGGCCAGCCTGCGGACACCAAAACTGACCTGGAGAGAGTAAAGATTTC	1555
OY	1517	TCAAGTGGACCATGTCATCTGGCCAGAT-GGCAACGAGTTGTCCTCTGCGCAGA- GGGT	1574
Db	1556	CCAATTTGATCATGTGATATGGCCTGATGGGCAAGAGATAGTATTGCTGGCAGAGGGGC	1615
OY	1575	CGTCTACTCAATTTGAGCTGCTCCACAGTTCACACCTTTTGTCTGTCCATCACAGCCACA	1634
Db	1616	CGTTTATGATCTCAGGTGCTCCACAGTGCCTACATTGTGCTCTCCATCACCTGCTAAC	1675
OY	1635	ACACAGGCTTTGGCACTGTAGAACTCTATAATGCAACCCGAGGGGGCATACAGCAGGAT	1694
Db	1676	ACTCAGGCTCTCGGCTTGTATGAACCTTTACAATGCTCTGAGGGGTGCTATTAAGCAGAT	1735
OY	1695	GTTGACTTGTCTCTAAGAAAATGGATGAATACGTTGCCAGCTTGCACTGCGCATCATTT	1754
Db	1736	GTTGACTTGTCTCTAAGAAAATGGATGAATGATGATGTGGCAAGCCTACACCTGCCCCACTTT	1795
OY	1755	GATGCCACACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAACAAAAAT	1814
Db	1796	GATGCCACACCTGACAGAGCTTAACAGATGAACAAGCCAAAGTATCTGGGACTCAACAAAGAT	1855
OY	1815	GGGCCATTCAAACCTAATTAATACAGATA 1843	
Db	1856	GGACCTTCAAGCCTTAAGCTACTACAGGTA 1884	

RESULT 2	AL527928	938 bp	mRNA	linear	EST 13-FEB-2001
LOCUS	AL527928	LTI_NFL003_NBC3	Homo sapiens	CDNA clone	CS0DC027YC19 5
DEFINITION	AL527928	LTI_NFL003_NBC3	Homo sapiens	CDNA clone	CS0DC027YC19 5
ACCESSION	AL527928	GI:12791421			
VERSION	AL527928.1				
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
AUTHORS	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.				
TITLE	Full-length CDNA libraries and normalization				
JOURNAL	Unpublished (2001)				
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.				
FEATURES	Location/Qualifiers				
source	1. 938 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="CS0DC027YC19" /clone_11b="LTI_NFL003_NBC3" /sex="male" /tissue_type="neuroblastoma cells" /lab_host="DH10B" /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand CDNA was primed with a NotI-oligo(dt) primer. Five prime end enriched, double-stranded CDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact: Feng Liang, Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email: fliang@lifetech.com URL : http://fulllength.invitrogen.com"				
BASE COUNT	241 a 208 c 275 g 212 t				
ORIGIN					
Query Match	50.2%; Score 926.2; DB 9; Length 938;				
Best Local Similarity	99.7%; Pred. No. 3.9e-224;				
Matches 936; Conservative	2; Mismatches 0; Indels 1; Gaps 1;				
QY	289 TCGGGGAGAGCTGAAGCAGGCGCCCAAGAGATCGAGAGCGCCGAGAGTACTCCTCATGG	348			
Db	1 TCGGGGAGAGCTGAAGCAGGCGCCCAAGAGATCGAGAGCGCCGAGAGTACTCCTCATGG	60			
QY	349 CCACCGTCACCAAGGCGCCCAAGAGCAATCCAGTTTGTCTGATGACATGCAGAGATTCA	408			
Db	61 CCACCGTCACCAAGGCGCCCAAGAGCAATCCAGTTTGTCTGATGACATGCAGAGATTCA	120			
QY	409 CCAAAATTCCTCCCAACCAAACTGGCCAGAGATCTTGTCTCGCTCGATCTCACAGTCTCCA	468			
Db	121 CCAAAATTCCTCCCAACCAAACTGGCCAGAGATCTTGTCTCGCTCGATCTCACAGTCTCCA	180			
QY	469 CTGACAGTACAGTTCAGCTGCATCTACACAGATAGCTCTGATGATGAGGTTTCTCCCC	528			
Db	181 CTGACAGTACAGTTCAGCTGCATCTACACAGATAGCTCTGATGATGAGGTTTCTCCCC	240			
QY	529 GAGAGAGCAGCAAAACCAACTCCAAAGGCGAGCAGCAATTTCTGTGTGAAGAACATCAAGC	588			
Db	241 GAGAGAGCAGCAAAACCAACTCCAAAGGCGAGCAGCAATTTCTGTGTGAAGAACATCAAGC	300			
QY	589 AGGCAAGATTGGACGCGCGGAGATTGAGATTGCAGAGCAAGACATGCTCTGATTT	648			
Db	301 AGGCAAGATTGGACGCGCGGAGATTGAGATTGCAGAGCAAGACATGCTCTGATTT	360			
QY	649 CACTCAGCAAAACGCTGCTCAGGGGGAGAGACCCCTTGGCTGCTGCTAAATAGTGGGCTGTA	708			
Db	361 CACTCAGCAAAACGCTGCTCAGGGGGAGAGACCCCTTGGCTGCTGCTAAATAGTGGGCTGTA	420			

QY 709 CACACATCACAGCCAGACAGCGGTGTGATGAGACACTCTGTGCCCTGGGGCTCAGT 768
 |||||||
 Db 421 CACACATCACAGCCAGACAGCGGTGTGATGAGACACTCTGTGCCCTGGGGCTCAGT 480

QY 769 GCCGCTGTCTGTGTACATCTACTCACTCAGATGAGTAGCTGCAGCACTGGCTG 828
 |||||||
 Db 481 GCCGCTGTCTGTGTWACATCTACTCACTCAGATGAGTAGCTGCAGCACTGGCTG 540

QY 829 AGCGTGAGTTCAGTCTTCGCTTGGAGAGGGCAGTCAGAGATGACTTCTGGTGTGA 888
 |||||||
 Db 541 AGCGTGAGTTCAGTCTTCGCTTGGAGAGGGCAGTCAGAGATGACTTCTGGTGTGA 600

QY 889 TTGACCGCTGTGTGACATGATGGGTGGCAGGCCAACATGATCCTGATGATGGGGAG 948
 |||||||
 Db 601 TTGACCGCTGTGTGACATGATGGGTGGCAGGCCAACATGATCCTGATGATGGGGAG 660

QY 949 ACTTAACCCACTGGGTTTATAGAAGTATCCAAAGCTGTTTAAAGAAGATCCGAGGCAATG 1008
 |||||||
 Db 661 ACTTAACCCACTGGGTTTATAGAAGTATCCAAAGCTGTTTAAAGAAGATCCGAGGCAATG 720

QY 1009 TGAAGAGAGCGTGTGCTGTTCACAGCGCTGTATCAGCTCTCCAAAGCTGGAGAGCTCT 1068
 |||||||
 Db 721 TGAAGAGAGCGTGTGCTGTTCACAGCGCTGTATCAGCTCTCCAAAGCTGGAGAGCTCT 780

QY 1069 GTGTCCCGCCATGAACGTCAATGATCTGTACCAACAGAGATTGATTAATCTGTACT 1128
 |||||||
 Db 781 GTGTCCCGCCATGAACGTCAATGATCTGTACCAACAGAGATTGATTAATCTGTACT 840

QY 1129 GCTGCCGAGATTCATTTGGATGGCCTGAAGAGACCACAGATGTGATGTTGGTGGGA 1188
 |||||||
 Db 841 GCTGCCGAGATTCATTTGGATGGCCTGAAGAGACCACAGATGTGATGTTGGTGGG- 899

QY 1189 AACAGTGTGTGTGTGCTATGTGAGGTAGGCAAGG 1227
 |||||||
 Db 900 AAMAAGTGTGTGTGTGCTATGTGAGGTAGGCAAGG 938

RESULT 3
 AL551097 921 bp mRNA linear EST 16-FEB-2001
 LOCUS AL551097 LTI_NFL006_PL2 Homo sapiens cDNA clone CSODI066YM07 5
 DEFINITION prime, mRNA sequence.
 ACCESSION AL551097
 VERSION AL551097.1 GI:12888715
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 921)
 AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 FEATURES
 source
 1. 921
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CSODI066YM07"
 /clone_1b="LTI_NFL006_PL2"
 /tissue_type="placenta"
 /note="Vector: PCMVSPORT 6; Site_1: NotI; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-stranded cDNA was digested with Not I and
 cloned into the Not I and Eco RV sites of the PCMVSPORT 6
 vector. Library was normalized. Library was constructed by
 Life Technologies. Contact : Peng Liang Life Technologies,
 a division of Invitrogen 9800 Medical Center Drive
 Rockville, Maryland 20850, USA Fax : (1) 301 610 8371

Email : fltangellifetech.com URL :
<http://fulllength.invitrogen.com>
 BASE COUNT 217 a 231 c 302 g 170 t 1 others
 ORIGIN

Query Match 48.7%; Score 897.4; DB 9; Length 921;
 Best Local Similarity 99.6%; Pred. No. 8e-217;
 Matches 909; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 69 GCCTTGGGCTGCCGACAGACAGAGCGGTGGGCCACAGCACTCAGAGAGCCGAGCAGCTC 128
 |||||||
 Db 1 GCCTTGGGCTGCCGACAGACAGAGCGGTGGGCCACAGCACTCAGAGAGCCGAGCAGCTC 60

QY 129 GACGAGGGGGCCGAGAGAGGGTGGCGATCGCTGTGAGAGGGCGCGCGCGAGCC 188
 |||||||
 Db 61 GACGAGGGGGCCGAGAGAGGGTGGCGATCGCTGTGAGAGGGCGCGCGCGAGCC 120

QY 189 GGGCGGGCCGAGAGGGGGAAAGAGCGGGGGGGGGTCAAGCCGTGGCGGGCGGGC 248
 |||||||
 Db 121 GGGCGGGCCGAGAGGGGGAAAGAGCGGGGGGGGGTCAAGCCGTGGCGGGCGGGC 180

QY 249 GGGGGAATGTGATGCCCTGACCGCGATGCCCGTGGCGGGGGTGGGAGAGCTGAACAG 308
 |||||||
 Db 181 GGGGGAATGTGATGCCCTGACCGCGATGCCCGTGGCGGGGGTGGGAGAGCTGAACAG 240

QY 309 GCCAAGGAGATCGAGAGCGCCGAGAGTACTCCTTCATGGCCACCGTCACCAAGGGCCC 368
 |||||||
 Db 241 GCCAAGGAGATCGAGAGCGCCGAGAGTACTCCTTCATGGCCACCGTCACCAAGGGCCC 300

QY 369 AAGAAGCAATCCAGTTTCTGTGATGACATGACAGAGTTCACCAATTCGCCACCAAACT 428
 |||||||
 Db 301 AAGAAGCAATCCAGTTTCTGTGATGACATGACAGAGTTCACCAATTCGCCACCAAACT 360

QY 429 GGCCGAAGATCTTGTCTCGCTCGATCTCACAGTCCCTCCACTGACAGCTACAGTTCAGCT 488
 |||||||
 Db 361 GGCCGAAGATCTTGTCTCGCTCGATCTCACAGTCCCTCCACTGACAGCTACAGTTCAGCT 420

QY 489 GCATCCTACACAGATAGCTCTGATGATGAGGTTTCTCCCGAGAGAGACGACCAACCAAC 548
 |||||||
 Db 421 GCATCCTACACAGATAGCTCTGATGATGAGGTTTCTCCCGAGAGAGACGACCAACCAAC 480

QY 549 TCCAAAGGCGAGCAATTTCTGTGTGAAGAACATCAAGCAGGCAATTTGGAGCCGG 608
 |||||||
 Db 481 TCCAAAGGCGAGCAATTTCTGTGTGAAGAACATCAAGCAGGCAATTTGGAGCCGG 540

QY 609 GAGATTGAGATTGACAGCAAGACATGCTGCTCTGATTTCACTCAGGAAACGTGCTCAG 668
 |||||||
 Db 541 GAGATTGAGATTGACAGCAAGACATGCTGCTCTGATTTCACTCAGGAAACGTGCTCAG 600

QY 669 GGGGAGAGCCCTTGGCTGGTGTAAATAGTGGGCTGTACACACATCACAGCCAGACA 728
 |||||||
 Db 601 GGGGAGAGCCCTTGGCTGGTGTAAATAGTGGGCTGTACACACATCACAGCCAGACA 660

QY 729 GCGGTGTGATTGAGACACTCTGTGCCCTGGGGCTCAGTGCCGCTGTGCTGTAAAC 788
 |||||||
 Db 661 GCGGTGTGATTGAGACACTCTGTGCCCTGGGGCTCAGTGCCGCTGTGCTGTAAAC 720

QY 789 ATCTACTCACTCAGATGAAGTAGCTGCAGACAGTGGCTGAGGCTGAGATTGACAGTTC 848
 |||||||
 Db 721 ATCTACTCACTCAGATGAAGTAGCTGCAGACAGTGGCTGAGGCTGAGATTGACAGTTC 780

QY 849 GCTTGAAGGGCGAGTCAGAGATGACTTCTGTGGTGTATTGACCGGCTGTGTGAACATG 908
 |||||||
 Db 781 GCTTGAAGGGCGAGTCAGAGATGACTTCTGTGGTGTATTGACCGGCTGTGTGAACATG 840

QY 909 GATGGGTGCAGGCCACATGATCCTGATGATGGGGAGACTTAACCCACTGGGTTAT 968
 |||||||
 Db 841 GATGGGTGCAGGCCACATGATCCTGATGATGGGGAGACTTAA-CCACTGGGTTAT 899

QY 969 AAGAAGTATCAA 981
 |||||||
 Db 900 AAGAAGTATCAA 912

RESULT 4
AL520704
LOCUS
DEFINITION AL520704 891 bp mRNA linear EST 13-FEB-2001
prime, mRNA sequence.
ACCESSION AL520704
VERSION AL520704.1 GI:12784197
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 891)
11, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1 . . 891

```

source
1. .891
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DB002YL22"
/clone_1lb="LTI_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : liang@lifetech.com
http://fulllength.invitrogen.com"
194 a 238 c 299 g 158. t 2 others
BASE COUNT
ORIGIN

```

Query Match	47.38;	Score 873;	DB 9;	Length 891;
Best Local Similarity	99.58;	Pred. No. 1.2e-210;		
Matches 873; Conservative	2;	Mismatches 2;	Indels 0;	Gaps 0;

QY	1	GGCGCGGGCAGGTCGGAGCTCGAGCTGCTGCTTCTGTGTTCTCTGTGGCCACCGTGCT	60
Db	15	GGCGCGGGCAGGTCGGAGCTCGAGCTGCTGCTTCTGTGTTCTGTGTGGCCCGCTGCT	74
QY	61	GTCCGGCTGCCCTGGGCTGCCGAACAGACAAGGCGTGGGCCACAGCAGCTCAGAAAGCCGA	120
Db	75	GTCCGGCTGCCCTGGGCTGCCGAACAGACAAGGCGTGGGCCACAGCAGCTCAGAAAGCCGA	134
QY	121	CGCAGCTCGACGCGAGGGGGCCGCGAGAGGGGTGGCGATCGCGTGTGGAAGGCGCGCGC	180
Db	135	CGCAGCTCGACGCGAGGGGGCCGCGAGAGGGGTGGCGATCGCGTGTGGAAGGCGCGCGC	194
QY	181	GGCGAGGCGGGCGGGCGCCAGAGGGGGAAAGAGGGGGGGCGGGGTCAAGCCGCTGGCC	240
Db	195	GGCGAGGCGGGCGGGCGCCAGARGGGAAARAAGCGGGGGGGGTCAAGCCGCTGGCC	254
QY	241	GGCGCGGGCGGGGGAATGTCATGCTTGAACCGCATGCCGCTGCCCGGGGGTGGGGAGGAGC	300
Db	255	GGCGCGGGCGGGGGAATGTCATGCTTGAACCGCATGCCGCTGCCCGGGGGTGGGGAGGAGC	314
QY	301	TGAAGCAGGCCAAGAGATCGAGACGCGGAGAGTACTCTTCATGGCCACCGTCACCA	360
Db	315	TGAAGCAGGCCAAGAGATCGAGACGCGGAGAGTACTCTTCATGGCCACCGTCACCA	374
QY	361	AGCGGCCAAGAAGCAATCCAGTTTGTGTATGACATGCAGGAGTTCAACCAATTCCCCA	420

Db	375	AGCGGCCCCAAGAAGCAAAATCCAGTTTGGCTGATGACATGCAGGAGTTCACCAATTTCGCCA	434
QY	421	CCAAAACCTGGCCCGAAGATCTTTGTCTGCTCGATCTCACAGTCTCTCCACTGCACAGCTACA	480
Db	435	CCAAAACCTGGCCCGAAGATCTTTGTCTGCTCGATCTCACAGTCTCTCCACTGCACAGCTACA	494
QY	481	GTTTCAGCTGCATCTCTACACAGATAGCTCTGATGATGAGGTTTCTCCCCGAGAGAAGCAGC	540
Db	495	GTTTCAGCTGCATCTCTACACAGATAGCTCTGATGATGAGGTTTCTCCCCGAGAGAAGCAGC	554
QY	541	AAACCAACTCCAAGGGCGACGACCAATTTCTGTGTGAAGACATCAAGCAGGCGAATTTG	600
Db	555	AAACCAACTCCAAGGGCGACGACCAATTTCTGTGTGAAGACATCAAGCAGGCGAATTTG	614
QY	601	GACGCGGGGAGATTGAGATTGCAGAGCAGACATGCTCTGCTCTGATTTTCACTCAGGAAC	660
Db	615	GACGCGGGGAGATTGAGATTGCAGAGCAGACATGCTCTGCTCTGATTTTCACTCAGGAAC	674
QY	661	GTGCTCAGGGGGAGAAAGCCCTTGCTGTGCTTAAATAGTGGGCTGTACACATCAG	720
Db	675	GTGCTCAGGGGGAGAAAGCCCTTGCTGTGCTTAAATAGTGGGCTGTACACATCAG	734
QY	721	CCCAAGACGCGGTGTGATTGAGACACTGTGTGCCCTGGGGGCTCAGTCCGCTGGTCTG	780
Db	735	CCCAAGACGCGGTGTGATTGAGACACTGTGTGCCCTGGGGGCTCAGTCCGCTGGTCTG	794
QY	781	CTTGTAACTCTACTCAACTCAGAATGAATGAGCTGCAGCACTGGCTGAGGCTGGAGTTG	840
Db	795	CTTGTAACTCTACTCAACTCAGAATGAATGAGCTGCAGCACTGGCTGAGGCTGGAGTTG	854
QY	841	CAGTGTTCGCTTGGAAAGGCGAGTCAGAAGATGACTT	877
Db	855	CAGTGTTCGCTTGGAAAGGCGAGTCAGAAGATGACTT	891

RESULT	5
BM449470	
LOCUS	
DEFINITION	1042 bp mRNA Homo sapiens CDNA clone IMAGE:5493677
AGENCOURT_6400881	
NIH_MGC_67	
Homo sapiens	
CDNA clone	
IMAGE:5493677	
EST	05-FEB-2002
BM449470	
5' mRNA sequence.	
BM449470	
BM449470.1	
GI:18498510	
FST	
KEYWORDS	
ACCESSION	
VERSION	

SOURCE	human.
ORGANISM	<i>Homo sapiens</i>
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 1042)
TITLE	NIH-MGC http://mgc.nci.nih.gov/ .
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished (1999)
	Contact: Robert Strausberg, Ph.D.

Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12117 row: m column: 06
High quality sequence stop: 694.

```

FEATURES
Source
Location/Qualifiers
1..1042
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5493677"
/clone_1lb="NIH_MGC_67"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life

```

BASE COUNT 267 a 219 c 309 g 240 t 7 others
ORIGIN

Query Match 46.8%; Score 862.6; DB 13; Length 1042;
Best Local Similarity 96.9%; Pred. No. 5.9e-208;
Matches 908; Conservative 0; Mismatches 25; Indels 4; Gaps 3;

532 AGAAGCAGCAAAACCACTCCAGGGCAGACAGCAATTTCTGTGAAGACATCAAGCAGG 591
1 AGAAGCAGCAAAACCACTCCAGGGCAGACAGCAATTTCTGTGAAGACATCAAGCAGG 60
592 CAGAATTGGACCGCCGGGAGATTGAGATTGCAGACCAAGACATGCTGCTGTGATTTCAC 651
61 CAGAATTGGACCGCCGGGAGATTGAGATTGCAGACCAAGACATGCTGCTGTGATTTCAC 120
652 TCAGAAACGCTGCTCAGGGGGAGAAAGCCCTGGCTGGTCTAAATAGTGGGCTGTACAC 711
121 TCAGAAACGCTGCTCAGGGGGAGAAAGCCCTGGCTGGTCTAAATAGTGGGCTGTACAC 180
712 ACATCAGACCCAGACAGCGGTGTGATTGAGACACTCTGCCCCGCGGCTCAGTGCC 771
181 ACATCAGACCCAGACAGCGGTGTGATTGAGACACTCTGCCCCGCGGCTCAGTGCC 240
772 GCTGCTGCTGTGAACATCTACTCAACTCAGAAATGAAGTAGCTGCAGACACTGGCTGAGG 831
241 GCTGCTGCTGTGAACATCTACTCAACTCAGAAATGAAGTAGCTGCAGACACTGGCTGAGG 300
832 CTGAGTGTGACGTGTCCTTGGAAGGGCGAGTCAGAAGATGACTTCTGCTGTATTTG 891
301 CTGAGTGTGACGTGTCCTTGGAAGGGCGAGTCAGAAGATGACTTCTGCTGTATTTG 360
892 ACCGCTGTGACATGATGGGTGGCAGGCCAATGATCTGTGATGATGGGGAGACT 951
361 ACCGCTGTGACATGATGGGTGGCAGGCCAATGATCTGTGATGATGGGGAGACT 420
952 TAACCCACTGGGTATTAAGAAGTATCCAAACGTTTAAAGAAGATCCGAGCATGTG 1011
421 TAACCCACTGGGTATTAAGAAGTATCCAAACGTTTAAAGAAGATCCGAGCATGTG 480
1012 AAGAGAGCGTGTGCTGTACAGGCTGTATCACTCTCCAAAGCTGGGAAGCTCTGTG 1071
481 AAGAGAGCGTGTGCTGTACAGGCTGTATCACTCTCCAAAGCTGGGAAGCTCTGTG 540
1072 TTCCGGCCATGAACGTCATGATTTCTGTACCAAAACAGAGTTGATTAATCTGACTGCT 1131
541 TTCCGGCCATGAACGTCATGATTTCTGTACCAAAACAGAGTTGATTAATCTGACTGCT 600
1132 GCCGAGATCCATTTGGATGGCCTGAAGAGACCAAGATGTGATGTTGGTGGGAAAC 1191
601 GCCGAGATCCATTTGGATGGCCTGAAGAGACCAAGATGTGATGTTGGTGGGAAAC 660
1192 AAGTGTGTGTGGCTATGTGAGGTAGGCAAGGGCTGCTGTGCTGTCTCAAAAGCTC 1251
661 AAGTGTGTGTGGCTATGTGAGGTAGGCAAGGGCTGCTGTGCTGTCTCAAAAGCTC 720
1252 TTGAGCAATTTGTACATTACCGAAATGACCCCATCTGTGCTGTGAGGCTGCATGG 1311
721 TTGAGCAATTTGTACATTACCGAAATGACCCCATCTGTGCTGTGAGGCTGCATGG 780
1312 ATGGG--TTACGGGTGCTAAAGCTAAATGAAGTCATCCGCAAGTCGATGCTAATAAC 1369
781 ATGGGTTACGGGTGCTAAAGCTAAATGAAGTCATCCGCAAGTCGATGCTAATAAC 840
1370 TTGCACAGGAATAGATAGTAGACACGGGAGCACTT-GGATCGCATGAATAACAGTT 1428
841 TTGCACAGGAATAGATAGTAGACACGGGAGCACTTGGGATCGCATGAATAACAGTT 900
1429 GTATC-GTATCATTATGGGCCACTCCACACAGAAA 1464
901 GTATCGGTATGCAATATGGGGCCANCTCAACCACANA 937

RESULT 6
AL519571 865 bp mRNA linear EST 13-FEB-2001
LOCUS
DEFINITION AL519571 LTI_NFL004_NBC2 Homo sapiens CDNA clone CS0DB004Y104 5
ACCESSION AL519571
VERSION AL519571
KEYWORDS prime, mRNA sequence.
SOURCE
ORGANISM human.
REFERENCE Homo sapiens
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL 1 (bases 1 to 865)
COMMENT Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source
1. 865
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DB004Y104"
/clone_1db="LTI_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: PCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the PCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"
BASE COUNT 193 a 231 c 284 g 155 t 2 others
ORIGIN

Query Match 46.4%; Score 856.2; DB 9; Length 865;
Best Local Similarity 99.2%; Pred. No. 2.2e-206;
Matches 858; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

8 GCAGTCGAGAGCTCGAGCTGCTGCTCTGCTCTCTGTCGACACCGTCGCTGCCGGC 67
1 GCAGTCGAGAGCTCGAGCTGCTGCTCTGCTCTCTGTCGACACCGTCGCTGCCGGC 60
68 TGCCTGGGCTGCCGAACAGCAAGGCGTGGGCCACAGCACTCAGAAAGCCGACAGCT 127
61 TGCCTGGGCTGCCGAACAGCAAGGCGTGGGCCACAGCACTCAGAAAGCCGACAGCT 120
128 CGACGAGGGGGCGGAGAGGGTGGCGATCGCTGTCGAGAGGGCGCGCGGGGAGG 187
121 CGACGAGGGGGCGGAGAGGGTGGCGATCGCTGTCGAGAGGGCGCGCGGGGAGG 180
188 CGGGCGGGCGCAGAGGGGGAAGAGCGGGGCGGGGCGGTCAGCCGCTGGCGGGCGG 247
181 CGGGCGGGCGCAGAGGGGGAAGAGCGGGGCGGGGCGGTCAGCCGCTGGCGGGCGG 240
248 CGGGGAGATGTCATGCTGACGCGATGCCGCTGCCGGGGGTGCGGGAGAGCTGAAGCA 307
241 CGGGGAGATGTCATGCTGACGCGATGCCGCTGCCGGGGGTGCGGGAGAGCTGAAGCA 300
308 GGGCAAGAGATCGAGAGCGCCGAGAGTACTCTTTCATGCGCCACCGTCACCAAGCGCC 367
301 GGGCAAGAGATCGAGAGCGCCGAGAGTACTCTTTCATGCGCCACCGTCACCAAGCGCC 360
368 CAGAAGCAATCCAGTTGCTGATGACATGACAGAGTTACCAAAATTTCCCAACCAAAAC 427
361 CAGAAGCAATCCAGTTGCTGATGACATGACAGAGTTACCAAAATTTCCCAACCAAAAC 420

QY 428 TGGCCGAGATCTTGTCTCGCTGATCTCAGCTCCCTCCACTGACAGCTACAGTTGAGC 487
|||||
Db 421 TGGCCGAGATCTTATCTCGCTGATCTCAGCTCCCTCCACTGACAGCTACAGTTGAGC 480
QY 488 TGCATCTACACAGATAGCTCTGATGATGAGGTTTCTCCCGAGAGAGAGCAACCA 547
|||||
Db 481 TGCATCTACACAGATAGCTCTGATGATGAGGTTTCTCCCGAGAGAGAGCAACCA 540
QY 548 CTCGAGGCGCAGCAGCAATTTCTGTGTGAAGACATCAAGCAGGAGAAATTGGAGCGCG 607
|||||
Db 541 CTCGAGGCGCAGCAGCAATTTCTGTGTGAAGACATCAAGCAGGAGAAATTGGAGCGCG 600
QY 608 GGAGATTGAGATTGCAGAGCAAGACATGTCTGCTGATTTCACTCAGGAAACGTGCTCA 667
|||||
Db 601 GGAGATTGAGATTGCAGAGCAAGACATGTCTGCTGATTTCACTCAGGAAACGTGCTCA 660
QY 668 GGGGAGAGAGCCCTTGGCTGTGCTAAATAGTGGGCTGTACACATCAAGCCAGAC 727
|||||
Db 661 GGGGAGAGAGCCCTTGGCTGTGCTAAATAGTGGGCTGTACACATCAAGCCAGAC 720
QY 728 AGCGGTGTGATTGAGACACTCTGTCCCTGGGGCTCAGTGCCTGTCTGTAA 787
|||||
Db 721 AGCGGTGTGATTGAGACACTCTGTCCCTGGGGCTCAGTGCCTGTCTGTAA 780
QY 788 CATCTACTCACTCAGAAATGAAGTAGCTGCAGCACTGGCTGAGGCTGAGTTGAGTGT 847
|||||
Db 781 CATCTACTCACTCAGAAATGAAGTAGCTGCAGCACTGGCTGAGGCTGAGTTGAGTGT 840
QY 848 CGCTTGAAGGGCGAGTCAGAGAT 872
|||||
Db 841 CGCTTGAAGGGCGAGTCAGAGAT 865

RESULT 7
BM468564 967 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT_6475681 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5578170
DEFINITION 5', mRNA sequence.

ACCESSION BM468564
VERSION BM468564.1 GI:18517606
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 967)
AUTHORS NIH-MGC http://mgi.ncl.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs.fda.gov
Tissue Procurement: ATCC/DCTD/DTF
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
http://image.llnl.gov
Plate: L1AM12332 row: m column: 19
High quality sequence stop: 685.

FEATURES
source location/Qualifiers
1. 967

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5578170"
/clone_id="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

BASE COUNT 214 a 253 c 328 g 172 t
ORIGIN
Query Match 45.9%; Score 847; DB 13; Length 967;
Best Local Similarity 94.4%; Pred. No. 5.2e-204;
Matches 900; Conservative 0; Mismatches 50; Indels 3; Gaps 2;

QY 1 GGGCGGGGCGAGCTGGAGCTGCTGCTCTGTTCTCTTGTGGCCACCGTGGCT 60
|||||
Db 3 GGGCGGGGCGAGCTGGAGCTGCTGCTCTGTTCTCTTGTGGCCACCGTGGCT 62
QY 61 GTCCGGCTGCTTGGGCTGCGGACAGAGCGGTGGCCACAGACCTCAGAGCCGA 120
|||||
Db 63 GTCCGGCTGCTTGGGCTGCGGACAGAGCGGTGGCCACAGACCTCAGAGCCGA 122
QY 121 CGCAGCTGAGAGCAGGGCGCGCAGAGAGGTTGGCGATCGCGTGTGGAAGCGCGCGC 180
|||||
Db 123 CGCAGCTGAGAGCAGGGCGCGCAGAGAGGTTGGCGATCGCGTGTGGAAGCGCGCGC 182
QY 181 GGGCAGGGCGGGCGCGCAGAGGGGGAAGAGCGGGGGCGGGGTCAGCCGTGGCC 240
|||||
Db 183 GGGCAGGGCGGGCGCGCAGAGGGGGAAGAGCGGGGGCGGGGTCAGCCGTGGCC 242
QY 241 GGGCGGGCGGGGGAATGTGATGCTGTACGCGATGCCGTGCCGGGGTGGGGAGAGC 300
|||||
Db 243 GGGCGGGCGGGGGAATGTGATGCTGTACGCGATGCCGTGCCGGGGTGGGGAGAGC 302
QY 301 TGAAGCAGGCCAAGAGATCGAGAGCGCGAGAGTACTCTTCATGCGCACCGTCACCA 360
|||||
Db 303 TGAAGCAGGCCAAGAGATCGAGAGCGCGAGAGTACTCTTCATGCGCACCGTCACCA 362
QY 361 AGCGGCCCAAGAGCAATCCAGTTGCTGATGACATGCAGAGTTCAACCAATTCCCA 420
|||||
Db 363 AGCGGCCCAAGAGCAATCCAGTTGCTGATGACATGCAGAGTTCAACCAATTCCCA 422
QY 421 CCAAACTGGCCGAAGATCTTGTCTGCTGATCTCAGATCTCAGAGCTCAGAGCTACA 480
|||||
Db 423 CCAAACTGGCCGAAGATCTTGTCTGCTGATCTCAGATCTCAGAGCTCAGAGCTACA 482
QY 481 GTTCACTGATCTCTACACAGATAGCTGTGATGATGAGGTTTCTCCCGAGAGAGCAGC 540
|||||
Db 483 GTTCACTGATCTCTACACAGATAGCTGTGATGATGAGGTTTCTCCCGAGAGAGCAGC 542
QY 541 AAACCACTCCAGAGGCGCAGAGCAATTTCTGTGTAAGAACATCAAGCAGAGATTTG 600
|||||
Db 543 AAACCACTCCAGAGGCGCAGAGCAATTTCTGTGTAAGAACATCAAGCAGAGATTTG 602
QY 601 GAGCGCGGAGATGAGATTGCAGAGCAAGACATGTCTGCTGATTTCACTCAGAGAAC 660
|||||
Db 603 GAGCGCGGAGATGAGATTGCAGAGCAAGACATGTCTGCTGATTTCACTCAGAGAAC 662
QY 661 GTGCTCAGGGGAGAGAGCCCTGCTGCTAAATAGTGGGCTGTACACACATCAGAG 720
|||||
Db 663 GTGCTCAGGGGAGAGAGCCCTGCTGCTAAATAGTGGGCTGTACACACATCAGAG 722
QY 721 CCCAGACAGCGGTGTTGATTGAGACACTGTGCTGCGGCTCAGTCCGCTGCTG 780
|||||
Db 723 CCCAGACAGCGGTGTTGATTGAGACACTGTGCTGCGGCTCAGTCCGCTGCTG 782
QY 781 CTGTAACTACTCACTCAGAAATGAAGTAGCTGCAGCACTGGCTGAGGCTGAGTTG 840
|||||
Db 783 CTGTAACTACTCACTCAGAAATGAAGTAGCTGCAGCACTGGCTGAGGCTGAGTTG 841
QY 841 CAGTGTGCTTGAAGGCGCAGTCAGAGATGAC--TTCTGTGTTGTTATGAGCGCTG 898
|||||
Db 842 CAGTGTGCTTGAAGGCGCAGTCAGAGATGACCTTCTGTTGTTGTTGAGCGCGC 901
QY 899 TGTGAACATGATGGTGGAGGCGCAACATGATCTGATGATGGGGAGACT 951
|||||
Db 902 TGGGTGAAATGATGGGGGGGCAAGGCCAACTGTGATCCCTGGAATGAAT 954

RESULT 8

AL554850
 LOCUS AL554850 858 bp mRNA linear EST 16-FEB-2001
 DEFINITION AL554850 LTI_NFL006_PL2 Homo sapiens cDNA clone CSODI087YL18 5
 prime, mRNA sequence.
 ACCESSION AL554850
 VERSION AL554850.1 GI:12896025
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 858)
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 FEATURES
 source
 1..858
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="LTI_NFL006_PL2"
 /tissue_type="placenta"
 /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
 was primed with a NotI-oligo(dt) primer. Five prime end
 enriched, double-stranded cDNA was digested with Not I and
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6
 vector. Library was normalized. Library was constructed by
 life technologies. Contact : Feng Liang Life Technologies,
 a division of Invitrogen 9800 Medical Center Drive
 Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
 Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com"

BASE COUNT 212 a 206 c 255 g 182 t 3 others
 ORIGIN

Query Match 45.9%; Score 845.8; DB 9; Length 858;
 Best Local Similarity 99.5%; Pred. No. 9.8e-204;
 Matches 855; Conservative 3; Mismatches 0; Indels 1; Gaps 1;

OY 233 CGCTGCGCGCGCGCGCGCGGGAATGTCATGCTGACGCGGATGCCGCTGCGCGCGGCGGTCGG 292
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1 CGCTGCGCGCGCGCGCGCGGGAATGTCATGCTGACGCGGATGCCGCTGCGCGCGGCGGTCGG 59
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 293 GGAGGAGCTGAAGCAGGCGCAAGGAGATGAGGAGCGCGGAGAGTACTCTTCATGGCCAC 352
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 60 GGAGGAGCTGAAGCAGGCGCAAGGAGATGAGGAGCGCGGAGAGTACTCTTCATGGCCAC 119
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 353 CGTACCAAGCGCGCGCGCGCGGGAATGTCATGCTGATGACATGCGAGAGTTCACCAA 412
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 120 CGTACCAAGCGCGCGCGCGGGAATGTCATGCTGATGACATGCGAGAGTTCACCAA 179
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 413 ATCCCAACCAAACTGGCGAAGATCTTGTCTCGCTCGATCTCAAGTCTCCACTGA 472
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 180 ATCCCAACCAAACTGGCGAAGATCTTGTCTCGCTCGATCTCAAGTCTCCACTGA 239
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 473 CAGCTACAGTTCAGCTGCATCTCTACAGATAGCTCTGATGATGAGGTTTCTCCCGAGA 532
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 240 CAGCTACAGTTCAGCTGCATCTCTACAGATAGCTCTGATGATGAGGTTTCTCCCGAGA 299
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 533 GAAGCAGCAACCAACTCCAAAGGGGAGCAGCAATTTCTGTGTGAAGACATCAAGCAGGC 592
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 300 GAAGCAGCAACCAACTCCAAAGGGGAGCAGCAATTTCTGTGTGAAGACATCAAGCAGGC 359
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 593 AGAATTGGACGCGCGGAGATGAGATTCAGAGCAAGACATGCTGCTGATTTCACT 652
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 360 AGAATTGGACGCGCGGAGATGAGATTCAGAGCAAGACATGCTGCTGATTTCACT 419
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 653 CAGGAACGTGCTCAGGGGAGAGACCCCTGGCTGCTAAATAGTGGGCTGTACACA 712
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 420 CAGGAACGTGCTCAGGGGAGAGACCCCTGGCTGCTAAATAGTGGGCTGTACACA 479
 OY 713 CATCACAGCCAGACAGCGGCTGTGATTTAGACACTCTGTGCCCTGGGGGCTAGTCCG 772
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 480 CATCACAGCCAGACAGCGGCTGTGATTTAGACACTCTGTGCCCTGGGGGCTAGTCCG 539
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 773 CTGCTGCTGTTAATCTACTCACTCAGATGAGTAGCTGCAGCAGTGGCTGAGGC 832
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 540 CTGCTGCTGTTAATCTACTCACTCAGATGAGTAGCTGCAGCAGTGGCTGAGGC 599
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 833 TGGAGTTGACAGTGTGCTGCTTGAAGGGCGAGTCAGAGATGACTTCTGCTGCTATTGA 892
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 600 TGGAGTTGACAGTGTGCTGCTTGAAGGGCGAGTCAGAGATGACTTCTGCTGCTATTGA 659
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 893 CCGCTGTGAACATGAGTGGGTGGCAGGCCAATGATCCTGATGATGGGGAGACTT 952
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 660 CCGCTGTGAACATGAGTGGGTGGCAGGCCAATGATCCTGATGATGGGGAGACTT 719
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 953 AACCCACTGGGTTTAAAGAGTATCCAAACGTTTAAAGAGATCCGAGCAATTGGA 1012
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 720 AACCCACTGGGTTTAAAGAGTATCCAAACGTTTAAAGAGATCCGAGCAATTGGA 779
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 1013 AGAGAGCGTACGTGCTTCACAGCGCTGTATCAGCTCTCCAAAGCTGGGAGCTGTGT 1072
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 780 AGAGAGCGTACGTGCTTCACAGCGCTGTATCAGCTCTCCAAAGCTGGGAGCTGTGT 839
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 1073 TCCGGCCATGACGTCAAT 1091
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 840 TCCGGCCATGACGTCAAT 858
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 9
 BM452485 1042 bp mRNA linear EST 05-FEB-2002
 LOCUS BM452485
 DEFINITION AGENCOURT_6400886 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5493438
 5', mRNA sequence.
 ACCESSION BM452485
 VERSION BM452485.1 GI:18501525
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1042)
 AUTHORS NIH-MGC http://mgc.ncl.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LAM12117 row: c column: 07
 High quality sequence stop: 667.

FEATURES
 source
 1..1042
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5493438"
 /clone_lib="NIH_MGC_67"
 /tissue_type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.75 kb. Library constructed by Life
 Technologies."

BASE COUNT 243 a 270 c 329 g 197 t 3 others
 ORIGIN

QY 609 GAGATTGAGATTGACAGACATGTCCTGATTTCACCTAGGAACGTCCTAG 668
|||||
DB 601 GAGATTGAGATTGACAGACATGTCCTGATTTCACCTAGGAACGTCCTAG 660
|||||
QY 669 GGGAGAGAGCCCTGCTGCTGCTAAATAGTGGCTGTACACACATCAGCCAGACA 728
|||||
DB 661 GGGAGAGAG-CCCTGCTGCTGCTAAATAGTGGCTGTACACACATCAGCCAGACA 719
|||||
QY 729 GCGGTGTTGATTGACACACTCTGTGCCCTGGGGCTCAGTGCCGCTGCTGTTGTAAC 788
|||||
DB 720 GCGGTGTTGATTGACACACTCTGTGCCCTGGGGCTCAGTGCCGCTGCTGTTGTAAC 779
|||||
QY 789 ATCTACTCACTCAGATGAAGTAGCTGCAGCAGCTGAGGCTGAGGCTGAGTGTGCTTC 848
|||||
DB 780 ATCTACTCACTCAGATGAAGTAGCTGCAGCAGCTGAGGCTGAGTGTGCTTC 839
|||||
RESULT 11
BM453920 867 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT_6402958 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5528995
DEFINITION 5', mRNA sequence.
ACCESSION BM453920
VERSION BM453920.1 GI:18502949
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 867)
AUTHORS NIH-MGC http://mgi.ncl.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCID/DTP
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM12206 row: 1 column: 20
High quality sequence stop: 670.
FEATURES
source
1.867
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5528995"
/clone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 222 a 201 c 241 g 202 t 1 others
ORIGIN
Query Match 43.8%; Score 807.4; DB 13; Length 867;
Best Local Similarity 97.7%; Pred. No. 5.6e-194;
Matches 840; Conservative 0; Mismatches 17; Indels 3; Gaps 2;

DB 121 GCTCGATCTCAGAGTCCCTCCACTGACAGCTACAGTTCAGCTGCATCTACACAGATAGCT 180
|||||
QY 508 CTGATGATGAGGTTCTCTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 567
|||||
DB 181 CTGATGATGAGGTTCTCTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
|||||
QY 568 TCTGTGTAAGAACATCAAGCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 627
|||||
DB 241 TCTGTGTAAGAACATCAAGCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
|||||
QY 628 AAGACATGTCCTGCTCTGATTCTCACTCAGGAAAGCTGCTCAGGGGAGAGAGAGAGAGAG 687
|||||
DB 301 AAGACATGTCCTGCTCTGATTCTCACTCAGGAAAGCTGCTCAGGGGAGAGAGAGAGAGAG 360
|||||
QY 688 GTGCTAAATAGTGGGCTGTACACACATCAGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAG 747
|||||
DB 361 GTGCTAAATAGTGGGCTGTACACACATCAGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAG 420
|||||
QY 748 TCTGTGCCCTGGGGCTCAGTGCCGCTGCTGCTGTTGTAACATCTACTCACTCAGAAATG 807
|||||
DB 421 TCTGTGCCCTGGGGCTCAGTGCCGCTGCTGCTGTTGTAACATCTACTCACTCAGAAATG 480
|||||
QY 808 AAGTAGCTCAGCAGCTGCTGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 867
|||||
DB 481 AAGTAGCTCAGCAGCTGCTGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 540
|||||
QY 868 AAGTAGCTTCTGTGCTGTATTGACCGCTGTGTAACATGAGGCTGAGGCTGAGGCTGAGGCT 927
|||||
DB 541 AAGTAGCTTCTGTGCTGTATTGACCGCTGTGTAACATGAGGCTGAGGCTGAGGCTGAGGCT 600
|||||
QY 928 TGATCCTGATGATGGGGAGAGACTTAACCCAGCTGGGTTTATAAGAGATATCCAACGTGT 987
|||||
DB 601 TGATCCTGATGATGGGGAGAGACTTAACCCAGCTGGGTTTATAAGAGATATCCAACGTGT 660
|||||
QY 988 TTAAGAGATCCGAGGATTTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1047
|||||
DB 661 TTAAGAGATCCGAGGATTTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
|||||
QY 1048 TCTCCAAAGCTGGAGAGCTCTGTGTCCGGCATGAGCTCAATGATTTGTTACCAAC 1107
|||||
DB 721 TCTCCAAAGCTGGAGAGCTCTGTGTCCGGCATGAGCTCAATGATTTGTTACCAAC 780
|||||
QY 1108 AGAG-TTTGAATACTGTACTGCTGCCGAGAAATCC--ATTTGATGGCTGAAGAGGA 1164
|||||
DB 781 AGAGTTTGAATACTGTACTGCTGCCGAGAAATCCATTTTGGATGGGCTGAAGAGGG 840
|||||
QY 1165 CCACAGATGTGATTTGCT 1184
|||||
DB 841 ACCCCANAAATGGTGATGTT 860
|||||
RESULT 12
BG708016 871 bp mRNA linear EST 07-MAY-2001
LOCUS BG708016
DEFINITION 602671341P1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4793770 5',
mRNA sequence.
ACCESSION BG708016
VERSION BG708016.1 GI:13984941
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 871)
AUTHORS NIH-MGC http://mgi.ncl.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
Toshiyuki and Piero Carninci (RIKEN)

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>

plate: LLAM10674 row: b column: 11
 High quality sequence stop: 839.

FEATURES

source

1. 871
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4793770"
 /clone_1lb="NIH_MGC_96"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: BluescriptR (modified
 pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
 size-selected for average insert size 2.3 kb and
 normalized to ROP 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carlncl, in preparation). Library
 constructed by M. Brownstein (NIMH/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 238 a 190 c 235 g 207 t 1 others
 ORIGIN

Query Match 42.1%; Score 775.6; DB 12; Length 871;
 Best Local Similarity 98.5%; Pred. No. 6.7e-186;
 Matches 803; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

OY 908 GGATGGGTGGCAGGCCAATGATCTGGATGATGGGGAGACTTAACCCACTGGGTTA 967
 DB 5 GGATGGGTGGCAGGCCAATGATCTGGATGATGGGGAGACTTAACCCACTGGGTTA 64
 OY 968 TAAGAAGTATCCAAACGTTTAAAGATCCGAGGCATTTGGAAGAGAGCGTGACTGG 1027
 DB 65 TAAGAAGTATCCAAACGTTTAAAGATCCGAGGCATTTGGAAGAGAGCGTGACTGG 124
 OY 1028 TGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTGTCGGCCATGAACGT 1087
 DB 125 TGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTGTCGGCCATGAACGT 184
 OY 1088 CAATGATCTCTTACCAACAGAGTTTGATACTGTACTGTCCCGAGAATCCATTTT 1147
 DB 185 CAATGATCTCTTACCAACAGAGTTTGATACTGTACTGTCCCGAGAATCCATTTT 244
 OY 1148 GGATGGCTGAAGAGGACACAGATGTGATGTTGGTGGGAACAAGTGTGTGTGG 1207
 DB 245 GGATGGCTGAAGAGGACACAGATGTGATGTTGGTGGGAACAAGTGTGTGTGG 304
 OY 1208 CTATGCTGAGGTAGGCAAGGGCTGTGCTCTCAAGCTCTTGGAGCAATTTGCTA 1267
 DB 305 CTATGCTGAGGTAGGCAAGGGCTGTGCTCTCAAGCTCTTGGAGCAATTTGCTA 364
 OY 1268 CATTAACGAAATCGACCCCATCTGCTCTGACAGGCTGCATGGAATGGGTTCAGGGTGT 1327
 DB 365 CATTAACGAAATCGACCCCATCTGCTCTGACAGGCTGCATGGAATGGGTTCAGGGTGT 424
 OY 1328 AAAGCTAAATGAGTCAATCCGGCAGTGCATGCTAATAACTTGCACAGAAATAAGAA 1387
 DB 425 AAAGCTAAATGAGTCAATCCGGCAGTGCATGCTAATAACTTGCACAGAAATAAGAA 484
 OY 1388 TGTAGTGACACGGGAGCAGCTTGTGATGCAATGAAACAGTGTATGCAATATGGG 1447
 DB 485 TGTAGTGACACGGGAGCAGCTTGTGATGCAATGAAACAGTGTATGCAATATGGG 544
 OY 1448 CCACCTCAACACAGAAATCGATGTGACAGCCCTCCGCACTCCGGAGCTGACGTGGAGCG 1507
 DB 545 CCACCTCAACACAGAAATCGATGTGACAGCCCTCCGCACTCCGGAGCTGACGTGGAGCG 604
 OY 1508 AGTACGTTCTCAGGTGACCATGTCTATCTGGCCAGATGGCAACGAGTTGTCTCTCTGGC 1567
 DB 604 AGTACGTTCTCAGGTGACCATGTCTATCTGGCCAGATGGCAACGAGTTGTCTCTCTGGC 1567

DB 605 AGTACGTTCTCAGGTGACCATGTCTATCTGGCCAGATGGCAACGAGTTGTCTCTCTGGC 664
 OY 1568 AGAGGTCGTCTACTCAATTTGAGCTGCTCCACAGTCCACACTTGTCTGTCCATCAC 1627
 DB 665 AGAGGTCGTCTACTCAATTTGAGCTGCTCCACAGTCCACACTTGTCTGTCCATCAC 724
 OY 1628 AGCCACACACAGGCTTTGGCACT-GATAGACTCTATATATGACCCGAGGGCGGATACA 1686
 DB 725 AGNCACACACAGGCTTTGGCACTGATAGACTCTATATATGACCCGAGGGCGGATACA 783
 OY 1687 AGCAGATGTGTACTTGTCTCTTAAGAAATGGAT 1721
 DB 784 AGCAGATGTGTACTTGTCTCTTAAGAAATGGAT 818

RESULT 13
 BQ641487 928 bp mRNA linear EST 15-JUL-2002
 LOCUS
 DEFINITION AGENCOURT_8290178 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:6305605
 5', mRNA sequence.
 ACCESSION BQ641487
 VERSION BQ641487.1 GI:21765659
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 928)
 AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LNCM2528 row: c column: 14
 High quality sequence stop: 663.

FEATURES
 source

1. 928
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6305605"
 /clone_1lb="NIH_MGC_43"
 /tissue_type="normal pigmented retinal epithelium"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pORF7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Library constructed by Ling Hong
 in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-CDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library. 1"

BASE COUNT 201 a 252 c 313 g 162 t

ORIGIN
 Query Match 41.9%; Score 773.4; DB 14; Length 928;
 Best Local Similarity 97.7%; Pred. No. 2.5e-185;
 Matches 825; Conservative 0; Mismatches 15; Indels 4; Gaps 4;

OY 2 GCGCGGCGAGTCGAGCTGCTGCTCTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCG 61
 DB 1 GCGCGGCGAGTCGAGCTGCTGCTCTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCG 60
 OY 62 TCCGGCTGCTTGGGCTGCCGACAGACAGAGCGGTGGCCACAGCAGCAGCAGCAGCAG 121
 DB 61 TCCGGCTGCTTGGGCTGCCGACAGACAGAGCGGTGGCCACAGCAGCAGCAGCAGCAG 120

```
QY 122 GCAGCTGACGACGAGCGCCGCGAGAGAGGTTGGCGATCGCGTGTGAGAGCGCCGCGC 181
    |||
Db 121 GCAGCTGACGACGAGCGCCGCGAGAGAGGTTGGCGATCGCGTGTGAGAGCGCCGCGC 180
QY 182 GGCAGCGCGCGCGCGCGCGAGAGCGGGAAGAGCGCGCGCGCGCGCGCGCGCGCGC 241
    |||
Db 181 GGCAGCGCGCGCGCGCGCGAGAGCGGGAAGAGCGCGCGCGCGCGCGCGCGCGCGC 240
QY 242 GCGCGCGCGCGCGCGCGATGTCATGCGCTGACCGGATGCCGCTGCCCGGGGCGGAGAGCT 301
    |||
Db 241 GCGCGCGCGCGCGCGCGATGTCATGCGCTGACCGGATGCCGCTGCCCGGGGCGGAGAGCT 300
QY 302 GAAGCAGCGCGCGAGATCGAGAGCGCGGAGAGTACTCTTCATGGCCACCGTCAACCA 361
    |||
Db 301 GAAGCAGCGCGCGAGATCGAGAGCGCGGAGAGTACTCTTCATGGCCACCGTCAACCA 360
QY 362 GCGCGCGCGAGCAAAATCCAGTTTGTGTGATGACATGACAGAGTTCAACCAATTCGCCAC 421
    |||
Db 361 GCGCGCGCGAGCAAAATCCAGTTTGTGTGATGACATGACAGAGTTCAACCAATTCGCCAC 420
QY 422 CAAACTGGCGGAGATCTTGTCTCGCTGATCTCAGAGTCTCCACTGACAGCTACAG 481
    |||
Db 421 CAAACTGGCGGAGATCTTGTCTCGCTGATCTCAGAGTCTCCACTGACAGCTACAG 480
QY 482 TTCAGCTGCATCTACACAGATAGCTGTGATGAGGTTTCTCCCGAGAGAGCAGCA 541
    |||
Db 481 TTCAGCTGCATCTACACAGATAGCTGTGATGAGGTTTCTCCCGAGAGAGCAGCA 540
QY 542 AACCACTCCAGAGCGCGAGCAATTTCTGTGTGAAGAATCAAGAGCAGAGATTGG 601
    |||
Db 541 AACCACTCCAGAGCGCGAGCAATTTCTGTGTGAAGAATCAAGAGCAGAGATTGG 600
QY 602 ACGCGGAGATGAGATTGACAGAGCAAGATGCTGCTGATTTCACTCAGGAACG 661
    |||
Db 601 ACGCGGAGATGAGATTGACAGAGCAAGATGCTGCTGATTTCACTCAGGAACG 660
QY 662 TGCTAGGGGAGAGAGCCCTTGCTGCTAAATAAGTGGCTGTACACATCACAGC 721
    |||
Db 661 TGCTAGGGGAGAGAGCCCTTGCTGCTAAATAAGTGGCTGTACACATCACAGC 720
QY 722 CCAGACAGCGGCTGTGATGAGACACTCTGT-GCCCTGGGGCTCAGT-GCCGCTGCT 779
    |||
Db 721 CCAGACAGCGGCTGTGATGAGACACTCTGTGCGCCCTGGGGCTCAATGGCCGCTGCT 780
QY 780 GCTGTAACTACTAC-TCAACTCAGAGTGAAGTACT-GCAGCACTGCTGAGGCTGAG 837
    |||
Db 781 GCTGTAACTACTACTTCCACTCAAAATGAAGTAACTGGCCGCACTGCTGAGGCTGAA 840
QY 838 TTGC 841
    ||
Db 841 TTTC 844
```

RESULT 14
BI668412 857 bp mRNA linear EST 12-SEP-2001
LOCUS 603292688F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5312133 5',
DEFINITION mRNA sequence.
ACCESSION BI668412
VERSION BI668412 GI:15582645
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 857)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shireki

```
Toshiyuki and Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLNL1790 row: p column: 22  
High quality sequence stop: 755.  
Location/Qualifiers  
1. 857  
/organism="Homo sapiens"  
/db_xref="taxon:9606"  
/clone="IMAGE:5312133"  
/clone_1lb="NIH_MGC_96"  
/issue_type="hypothesis"  
/lab_host="DH10B"  
/note="Organ: brain; Vector: pBluescript (modified  
pBluescript KS+); Site: 1: BamHI; Site: 2: SalI-XhoI (gtcgaag  
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',  
size-selected for average insert size 2.3 kb and  
normalized to 10^5. This is a primary library enriched  
for full-length clones and constructed using the  
Cap-trapper method (Carninci, in preparation). Library  
constructed by M. Brownstein (NHGRI, National  
Institutes of Health). Note: this is a NIH_MGC Library."  
BASE COUNT 219 a 208 c 241 g 189 t  
ORIGIN  
Query Match 41.3%; Score 762.2; DB 13; Length 857;  
Best Local Similarity 96.3%; Pred. No. 1.7e-182;  
Matches 823; Conservative 0; Mismatches 28; Indels 4; Gaps 4;  
QY 291 GGGAGAGCTGAGCAG-GCCAAGAGATCGAGAGCGCCGAGAGTACTCTTCATGCG 349  
    |||  
Db 2 GCGGAGAGCTGAGCAGTGCACAGAGATCGAGAGCGCCGAGAGTACTCTTCATGCG 61  
QY 350 CACCGTACCAAGCGCCCAAGAGCAAAATCCAGTTTGTGATGACATGAGAGTTCA 409  
    |||  
Db 62 CACCGTACCAAGCGCCCAAGAGCAAAATCCAGTTTGTGATGACATGAGAGTTCA 121  
QY 410 CAAATCCCCCAAAACTGGCCGA-AGATCTTGTCTGCTGATCTCAGTCTCCTCA 468  
    |||  
Db 122 CAAATCCCCCAAAACTGGCCGAAGATCTTGTCTGCTGATCTCAGTCTCCTCA 181  
QY 469 CTGACAGCTACAGTTCAAGTCTGATCTCAGATAGTCTGATGATGAGTTTCTCC 528  
    |||  
Db 182 CTGACAGCTACAGTTCAAGTCTGATCTCAGATAGTCTGATGATGAGTTTCTCC 241  
QY 529 GAGAGAAGCAAACTCAAGGCGAGCAATTTCTGTGAAGAACATCAAGC 588  
    |||  
Db 242 GAGAGAAGCAAACTCAAGGCGAGCAATTTCTGTGAAGAACATCAAGC 301  
QY 589 AGCAGATTTGAGCGCGGAGATTGAGATGAGAGCAAGACATGCTCTGATT 648  
    |||  
Db 302 AGCAGATTTGAGCGCGGAGATTGAGATGAGAGCAAGACATGCTCTGATT 361  
QY 649 CACTCAGGAAGCTGCTCAGGGGAGAGAGCCCTGCTGCTAAATAGTGGCTGTA 708  
    |||  
Db 362 CACTCAGGAAGCTGCTCAGGGGAGAGAGCCCTGCTGCTAAATAGTGGCTGTA 421  
QY 709 CACACATCACAGCCAGACAGCGGCTGTGATGAGACACTCTGTGCCCTGGGGCTCAGT 768  
    |||  
Db 422 CACACATCACAGCCAGACAGCGGCTGTGATGAGACACTCTGTGCCCTGGGGCTCAGT 481  
QY 769 GCGCGTGTCTGCTGTACATCTACTCACTCAGATGAAGTAGTGCAGCACTGGCTG 828  
    |||  
Db 482 GCGCGTGTCTGCTGTACATCTACTCACTCAGATGAAGTAGTGCAGCACTGGCTG 541  
QY 829 AGGCTGAGTTGAGTGTCTGCTTGAAGGCGAGTCAAGAGATGACTTCTGCTGTA 888  
    |||  
Db 542 AGGCTGAGTTGAGTGTCTGCTTGAAGGCGAGTCAAGAGATGACTTCTGCTGTA 601  
QY 889 TTGACCGCTGTGTAACATGATGGGTGGAGCCCAACATGATCTGATGAGGGAG 948
```

Db	602	TTGACCGCTGTGTGAACATGGATGGGTGGCAGGCCAACATGATCCTGGATGATGGGGAG	661
QY	949	ACTT-AAACCACCTGGGTTTATAAGAAGTATCCAAACGTGTTTAAGAAGAT-CCGAGGCAT	1006
Db	662	ACTTAACCTACTGGGTATATAAAAGTTTCCCAACGTGTTTAAGAAGATCCCGAGGCAT	721
QY	1007	TGTGGAAGAGAGCGCTGACTGCTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCT	1066
Db	722	TGTGGAAGAGAGAGCGGTGACTGCTGTTCACAGGCTGTATCAGCTCTCCGAAGCTGGGAAGCT	781
QY	1067	CTGTGTTCCGGCCATGAACGTCAATGATTCTGTATCCCAACAGAAGTTGATTAACCTTGTA	1126
Db	782	CTGTGTTCCGGCCCTGAACGTCACTAATTCTGTATCCCAACCGGAAGTCGCATACCTGTTC	841
QY	1127	CTGCTGCCGAGAATC	1141
Db	842	ATGCTGCCGGGAATC	856

RESULT	15				
LOCUS	AU126038				
DEFINITION	AU126038	764 bp	mRNA	linear	EST 01-AUG-2002
	AU126038	NT2RM4	Homo sapiens	CDNA clone	NT2RM4002613 5', mRNA
ACCESSION	AU126038				
VERSION	AU126038.1	GI:10950754			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 764)				
AUTHORS	Ota, T., Wakamatsu, A., Ozawa, M., Ishii, S., Saito, K., Yamamoto, J., Nakamura, Y., Nishikawa, T., Nagai, T., Suzuki, Y., Sugano, S. and Isogai, T.				
TITLE	HRI human CDNA project (Ota, T., Wakamatsu, A., Ozawa, M., Ishii, S., Saito, K., Yamamoto, J., Nakamura, Y., Nishikawa, T., Nagai, T., Suzuki, Y., Sugano, S., Isogai, T.)				
JOURNAL	Unpublished (2000)				
COMMENT	Contact: Takao Isogai				

Unpublished (2000)
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human CDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.

FEATURES	Location/Qualifiers
Source	1. .764

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RM4002613"
/clone_1b="NT2RM4"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/note="Vector: pME18SFL3; mRNA from uninduced NT2 neuronal precursor cells"

```

BASE COUNT	193 a	157 c	228 g	183 t	3 others
ORIGIN					
Query Match		39.78;	Score 732.8;	DB 9;	length 764;
Best Local Similarity		99.1%;	Pred. No. 4.6e-175;		
Matches 756; Conservative		0;	Mismatches 5;	Indels 2;	Gaps 2;

QY	472	ACAGCTACAGTTCAGCTGCATCCTACACACAGATAGCTCTGATGATGAGGTTTCTCCCCGAG	531
Db	1	ACAGCTACAGTTCAGCTGCATCCTACACACAGATAGCTCTGATGATGAGGTTTCTCCCCGAG	60
QY	532	AGAGGACGCAACCACTCCAGGGCAGCAGCAATTCTGTGTGAAGACATCAAGCAGG	591

Db	61	AGAGCAGCAAAACCACTCCAGGGCAGCAGCAATTTCTGTGAGAGACATCAAGCAGG	120
QY	592	CAGAAATTTGGACGCCGGGAGATTGAGATTGCAGAGCAAGACATGTCCTCTGATTTCAC	651
Db	121	CAGAAATTTGGACGCCGGGAGATTGAGATTGCAGAGCAAGACATGTCCTCTGATTTCAC	180
QY	652	TCAGGAAACGTGCTCAGGGGGGAGAAAGCCCTTGGCTGTGCTAAATAGTGGGCTGTACAC	711
Db	181	TCAGGAAACGTGCTCAGGGGGGAGAAAGCCCTTGGCTGTGCTAAATAGTGGGCTGTACAC	240
QY	712	ACATCACAGCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGGGGGCTCAGTGC	771
Db	241	ACATCACAGCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGGGGGCTCAGTGC	300
QY	772	GCTGCTCTGCTTGTAAACATCTACTCACTCAGAATGAAGTAGCTGCAGCACTGGCTGAGG	831
Db	301	GCTGCTCTGCTTGTAAACATCTACTCACTCAGAATGAAGTAGCTGCAGCACTGGCTGAGG	360
QY	832	CTGAGATTGCAGTGTTCGCTTGGAAGGGCGAGTCAGAAGATGACTTCTGGTGTGATTG	891
Db	361	CTGAGATTGCAGTGTTCGCTTGGAAGGGCGAGTCAGAAGATGACTTCTGGTGTGATTG	420
QY	892	ACCGCTGTGTGAACATGATGGGTGGCAGGCCAACATGATCTCGATGATGGGGGAGACT	951
Db	421	ACCGCTGTGTGAACATGATGGGTGGCAGGCCAACATGATCTCGATGATGGGGGAGACT	480
QY	952	TAAACCACTGGGTTTATTAAGAGATATCCAACGTGTTTAAAGAGATCCGAGCATTTGG	1011
Db	481	TAAACCACTGGGTTTATTAAGAGATATCCAACGTGTTTAAAGAGATCCGAGCATTTGG	540
QY	1012	AAGAGAGCGTGAAGTGTGTTCAAGCGCTGTATCAGCTCTCCAAAAGCTGGGAAGCTGTG	1071
Db	541	AAGAGAGCGTGAAGTGTGTTCAAGCGCTGTATCAGCTCTCCAAAAGCTGGGAAGCTGTG	600
QY	1072	TTCGGGCCATGAAGCTCAATGATTTCTGTTACCAAAACAGAGTTTGATACTTGTACTGCT	1131
Db	601	TTCGGGCCATGAAGCTCAATGATTTCTGTTACCAAAACAGAGTTTGAT- ACTTGTACTGCT	659
QY	1132	GCCGAGAATCCATTTTGGATGGCCTGAAGAGGACACAGATGTGATGTTGGTGGGAAC	1191
Db	660	GCCGAGAATCCATTTTGGATGGCCTGAANAGGACACAGATGTGATGTTGGTGGGAAC	719
QY	1192	AAGTGTGTGTGTGCTATGCTG- AGGTAGGCAAGGGCTGCT 1233	
Db	720	AAGTGTGTGTGTGCTATGCTGAAGGTANGCCANGGCTTCT 762	

Search completed: April 21, 2003, 03:59:15
Job time : 2497.43 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2003, 21:15:39 ; Search time 391.921 Seconds
(without alignments)
10595.712 Million cell updates/sec

Title: US-09-782-051-1_COPY_1_1844
Perfect score: 1844
Sequence: 1 ggcgcggcgcagtcgcgagct.....aacctaattaccagatac 1844

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1844	100.0	2563	19	AAV28617
2	1840.8	99.8	3634	23	ABV23195
3	1840.8	99.8	3634	23	ABV29032
4	1551.8	84.2	3616	20	AAV73924
5	950	51.5	5030	23	AAV76216
6	702.4	38.1	721	21	AAF16004
7	694.2	37.6	1566	23	ABL03693
8	649.2	35.2	1776	23	ABL29787
9	486.2	26.4	4406	23	ABL03692

C	10	461.4	25.0	553	24	ABL79045
C	11	422	22.9	504	23	ABV44141
C	12	421.6	22.9	5537	23	ABL29786
C	13	406.4	22.0	636	21	AA280766
C	14	402.4	21.8	2200	21	AAV7823
C	15	402.4	21.8	2658	23	ABV23126
C	16	402.4	21.8	2658	23	ABV28967
C	17	378.2	20.5	1592	23	ABL28157
C	18	363.8	19.7	435	23	ABV14215
C	19	337	18.3	384	24	ABN93848
C	20	320.4	17.4	349	21	AAA44806
C	21	314.4	17.0	2211	22	AAV91228
C	22	313.2	17.0	4380	23	ABL28156
C	23	245	13.3	939	23	AAV76217
C	24	244.8	13.3	312	23	ABV35309
C	25	242.8	13.2	1727	21	AAV46315
C	26	242.6	13.2	278	24	ABN73754
C	27	241.6	13.1	1461	22	AAV09278
C	28	236.2	12.8	1488	21	AAV39567
C	29	236.2	12.8	441529	22	AAV99682
C	30	233.6	12.7	1461	22	AAV09277
C	31	231.8	12.6	1728	21	AAV39836
C	32	228.8	12.4	1865	21	AAV44037
C	33	226.4	12.3	1494	21	AAV46817
C	34	225.6	12.2	1422	22	AAV65801
C	35	225.6	12.2	1557	22	AAV71862
C	36	225.6	12.2	1557	23	AAV96121
C	37	225.6	12.2	1939	24	AAV52394
C	38	225.6	12.2	349980	22	AAV68526
C	39	221	12.0	1812	17	AAV29194
C	40	217.8	11.8	1812	13	AAV29419
C	41	214.6	11.6	1767	17	AAV44513
C	42	208.2	11.3	58909	22	AAV28543
C	43	202	11.0	656	21	AAV14585
C	44	190	10.3	1877	18	AAV06354
C	45	189.2	10.3	1941	17	AAV44515

ALIGNMENTS

RESULT 1	AAV28617	standard; cDNA; 2563 BP.
ID	AAV28617	
XX	AAV28617	
AC	AAV28617	
XX	21-AUG-1998	(first entry)
DT		
DE		
XX		
XX		
KW	DDab9.3 enzyme; S-adenosyl-L-homocysteine hydrolase-type activity;	
KW	AHCY-type activity; stimulation; inhibition; dendritic cell; cancer;	
KW	autoimmune disease; transplantation; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	3..1847
FT		/*tag= a
FT		/product= "DD4b5.3 enzyme"
XX		
PN	W09814562-AL	
XX		
PD	09-APR-1998.	
XX		
XX	06-OCT-1997;	97WO-NZ00133.
PF		
XX		
PR	04-OCT-1996;	96NZ-0299507.
XX		
PA	(HART/) HART D N J.	
XX		
PI	Hart DNG;	

Human ovarian can
Human prostate exp
Drosophila melanog
Human colon cancer
Human cancer assoc
Human prostate exp
Human prostate exp
Drosophila melanog
Human prostate exp
Gene #346 used to
Human secreted exp
Human S-adenosylho
Drosophila melanog
DNA encoding novel
Human prostate exp
Arabidopsis thalia
Bovine embryonic g
Trichomonas vagina
M. tuberculosis an
Mycobacterium tube
Trichomonas vagina
Arabidopsis thalia
Zea mays DNA fragm
Arabidopsis thalia
C glutamicum codin
Corynebacterium gl
C. glutamicum gene
SahH gene encoding
C glutamicum codin
S-adenosylhomocyst
Nicotiana tabacum
Asparagus S-adenos
Genomic fragment #
Aspergillus oryzae
AL-2-long (AL-21)
Arabidopsis S-aden

OY 1201 TGTGTGCTATGTTGAGTAGGCAAGGGCTGTGTCTCTCTCAAGCTCTTGAGCAA 1260
DB 1271 TGTGTGCTATGTTGAGTAGGCAAGGGCTGTGTCTCTCTCAAGCTCTTGAGCAA 1330
OY 1261 TGTCTTACATTACCGAAATCGACCCCATCTGTGTCTCTGAGCCCTGCATGATGGGTCA 1320
DB 1331 TGTCTTACATTACCGAAATCGACCCCATCTGTGTCTCTGAGCCCTGCATGATGGGTCA 1390
OY 1321 GGGTGGTAAAGCTAAATGAGTATCTCCGGCAGTGTCTGTAATACTTGCACAGCAA 1380
DB 1391 GGGTGGTAAAGCTAAATGAGTATCTCCGGCAGTGTCTGTAATACTTGCACAGCAA 1450
OY 1381 ATAGATGTAGTACACGGAGCAGCTGTGATGCGATGAAACAGTTGATCGATGCA 1440
DB 1451 ATAGATGTAGTACACGGAGCAGCTGTGATGCGATGAAACAGTTGATCGATGCA 1510
OY 1441 ATATGGCCACTCCACACAGAAATCGATGTGACACAGCTCCGCACTCCGAGCTGAGT 1500
DB 1511 ATATGGCCACTCCACACAGAAATCGATGTGACACAGCTCCGCACTCCGAGCTGAGT 1570
OY 1501 GGGAGCGAGTACGTTCTCAGGTGGACCATGTCTGTGCGCAGATGGCAACGAGTTC 1560
DB 1571 GGGAGCGAGTACGTTCTCAGGTGGACCATGTCTGTGCGCAGATGGCAACGAGTTC 1630
OY 1561 TCCTGGCAGAGGCTCTCTACTCAATTGAGCTGCTCCACAGTCCACCTTTGTCTGT 1620
DB 1631 TCCTGGCAGAGGCTCTCTACTCAATTGAGCTGCTCCACAGTCCACCTTTGTCTGT 1690
OY 1621 CCATCACAGCCACACACAGGCTTTGGCAGCTGATAGACTCTATAATGACCCGAGGGC 1680
DB 1691 CCATCACAGCCACACACAGGCTTTGGCAGCTGATAGACTCTATAATGACCCGAGGGC 1750
OY 1681 GATACAAAGCAGGATGTGTACTTCTCTAGAAATGATGAATACGTTGCCAGCTGC 1740
DB 1751 GATACAAAGCAGGATGTGTACTTCTCTAGAAATGATGAATACGTTGCCAGCTGC 1810
OY 1741 ATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1800
DB 1811 ATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1870
OY 1801 GACTCAACAAAATGGGCCATTCAAACCTAATTATACAGATAC 1844
DB 1871 GACTCAACAAAATGGGCCATTCAAACCTAATTATACAGATAC 1914
RESULT 3
ABV29032
ID ABV29032 standard; cDNA; 3634 BP.
XX
AC ABV29032;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 29023.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Endege WO, Monahan JE;
PI WPI; 2001-662795/76.
XX
DR
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX
PS Claim 1; Page 6147-6148; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 3634 BP; 876 A; 966 C; 869 G; 918 T; 5 other;
Query Match 99.8%; Score 1840.8; DB 23; Length 3634;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1842; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 GGGCGGGCAGGTGCGAGCTCGAGCTGCTGCTTCTGTCTGTGGCCACCGTGCCT 60
DB 71 GGGCGGGCAGGTGCGAGCTCGAGCTGCTGCTTCTGTGTGGCCACCGTGCCT 130
OY 61 GTCCGCTGCTTGGCTGCCAACAAGAGCGGTGGCCACAGCAGCTCAGAGCCGA 120
DB 131 GTCCGCTGCTTGGCTGCCAACAAGAGCGGTGGCCACAGCAGCTCAGAGCCGA 190
OY 121 CGCAGCTGACGAGGGGGCCGAGAGGGGTGGCGATCGGTGCGAGGGCCGCGGC 180
DB 191 CGCAGCTGACGAGGGGGCCGAGAGGGGTGGCGATCGGTGCGAGGGCCGCGGC 250
OY 181 GGGCAGGGGGGGCCGAGAGGGGGAGAGGGGGGGGGGGGGGTCAGCCGCTGCC 240
DB 251 GGGCAGGGGGGGCCGAGAGGGGGAGAGGGGGGGGGGGGGGTCAGCCGCTGCC 310
OY 241 GGGCGGGCGGGGAATGTGATGCTGACGCGATGCCGCTGCCCGGGGTGGAGAGGC 300
DB 311 GGGCGGGCGGGGAATGTGATGCTGACGCGATGCCGCTGCCCGGGGTGGAGAGGC 370
OY 301 TGAAGCAGGCCAAGAGATGACGAGCGCCGAGAGTACTCTTCATGGCCACCGTCACA 360
DB 371 TGAAGCAGGCCAAGAGATGACGAGCGCCGAGAGTACTCTTCATGGCCACCGTCACA 430
OY 361 AGCGCCCCAAGAGCAATTCAGTTGCTGATGACATGACAGAGTTCACCAATTCGCCA 420
DB 431 AGCGCCCCAAGAGCAATTCAGTTGCTGATGACATGACAGAGTTCACCAATTCGCCA 490
OY 421 CCAAACTGGCCGAGATCTTGTCTGCTGCTGATCTCAGCTCTCCACTGACAGTACA 480
DB 491 CCAAACTGGCCGAGATCTTGTCTGCTGCTGATCTCAGCTCTCCACTGACAGTACA 550
OY 481 GTTCAGCTGATCTTACAGATAGCTCTGATGATGAGGTTCTCCCGAGAGAGCAGC 540
DB 551 GTTCAGCTGATCTTACAGATAGCTCTGATGATGAGGTTCTCCCGAGAGAGCAGC 610
OY 541 AAACCAACTCCAAAGGCGAGCAATTTCTGTGTGAAGACATCAAGCAGCAGAAATTG 600
DB 611 AAACCAACTCCAAAGGCGAGCAATTTCTGTGTGAAGACATCAAGCAGCAGAAATTG 670

```
OY 601 GACGCCGGAGATTGAGATGTCAGAGCAAGACATGTCGCTCTGATTTCACCTCAGGAAC 660
    |||
Db 671 GACGCCGGAGATTGAGATGTCAGAGCAAGACATGTCGCTCTGATTTCACCTCAGGAAC 730
OY 661 GTGCTCAGGGGAGAAAGCCCTGGCTGGTGTCTAAATAGTGGGCTGTACACACATCAG 720
    |||
Db 731 GTGCTCAGGGGAGAAAGCCCTGGCTGGTGTCTAAATAGTGGGCTGTACACACATCAG 790
OY 721 CCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCCTGGGGCTCAGTGGCGTGTCTG 780
    |||
Db 791 CCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCCTGGGGCTCAGTGGCGTGTCTG 850
OY 781 CTGTACATCTACTCACTCAGATGAAGTAGCTGCAGCACTGGCTGAGGCTGAGATTG 840
    |||
Db 851 CTGTACATCTACTCACTCAGATGAAGTAGCTGCAGCACTGGCTGAGGCTGAGATTG 910
OY 841 CAGTGTCCGCTTGAAGGGGCGAGTCAAGATGACTTCTGTGTGTATTGACCCGCTGTG 900
    |||
Db 911 CAGTGTCCGCTTGAAGGGGCGAGTCAAGATGACTTCTGTGTGTATTGACCCGCTGTG 970
OY 901 TGAACATGATGGGTGGCGAGGCCAACATGATCCTGGATGATGGGGGAGACTTAACCCACT 960
    |||
Db 971 TGAACATGATGGGTGGCGAGGCCAACATGATCCTGGATGATGGGGGAGACTTAACCCACT 1030
OY 961 GGGTTTATAGAAGTATCCAAACGTTTAAAGAAGATCCGAGCATTTGTGAAGAGAGCG 1020
    |||
Db 1031 GGGTTTATAGAAGTATCCAAACGTTTAAAGAAGATCCGAGCATTTGTGAAGAGAGCG 1090
OY 1021 TGACTGGTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAGCTCTGTGTCGGCCA 1080
    |||
Db 1091 TGACTGGTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAGCTCTGTGTCGGCCA 1150
OY 1081 TGAACGTCAATGATCTCTGTACCAACAAGAAGTTGATACTGTACTGCTGCCGAGAAT 1140
    |||
Db 1151 TGAACGTCAATGATCTCTGTACCAACAAGAAGTTGATACTGTACTGCTGCCGAGAAT 1210
OY 1141 CCATTTTGGATGGCCTGAAGAGACCACAGATGTGATGTTGGTGGGAACAAGTGTGG 1200
    |||
Db 1211 CCATTTTGGATGGCCTGAAGAGACCACAGATGTGATGTTGGTGGGAACAAGTGTGG 1270
OY 1201 TGTGTGGCTATGGTGAAGGTAGGCAAGGGCTGCTGTGCTCTCAAAAGCTCTTGAGCAA 1260
    |||
Db 1271 TGTGTGGCTATGGTGAAGGTAGGCAAGGGCTGCTGTGCTCTCAAAAGCTCTTGAGCAA 1330
OY 1261 TTGCTTACATTAACGGAATCGACCCCATCTGTGCTCTGACGGCTGCATGGATGGTTCA 1320
    |||
Db 1331 TTGCTTACATTAACGGAATCGACCCCATCTGTGCTCTGACGGCTGCATGGATGGTTCA 1390
OY 1321 GGGTGGTAAAGCTAAATGAAGTCAATCCGCAAGTCGATGCTGTAATACTTGACAGGAA 1380
    |||
Db 1391 GGGTGGTAAAGCTAAATGAAGTCAATCCGCAAGTCGATGCTGTAATACTTGACAGGAA 1450
OY 1381 ATAAGATGTAGTGACACGGGAGCACTTGATCGCATGAAAACAGTTGTATCGTATGCA 1440
    |||
Db 1451 ATAAGATGTAGTGACACGGGAGCACTTGATCGCATGAAAACAGTTGTATCGTATGCA 1510
OY 1441 ATATGGGCCACTCCAACACAGAAATCGATGTGACACAGCCCTCCGCACTCCGGAGCTGACGT 1500
    |||
Db 1511 ATATGGGCCACTCCAACACAGAAATCGATGTGACACAGCCCTCCGCACTCCGGAGCTGACGT 1570
OY 1501 GGGAGGAGTACGTTCTCAGGTGACCATGTCTATGCGCAGATGGCAAAAGAGTTGTCC 1560
    |||
Db 1571 GGGAGGAGTACGTTCTCAGGTGACCATGTCTATGCGCAGATGGCAAAAGAGTTGTCC 1630
OY 1561 TCCTGGCAGAGGGTGTCTACTCAATTTGAGCTGCTCCACAGTTCACCTTGTCTGT 1620
    |||
Db 1631 TCCTGGCAGAGGGTGTCTACTCAATTTGAGCTGCTCCACAGTTCACCTTGTCTGT 1690
OY 1621 CCATCAGACCAACAACAGAGGCTTTGGCACTGATAGAACTCTATAATGCAACCCGAGGGGC 1680
    |||
Db 1691 CCATCAGACCAACAACAGAGGCTTTGGCACTGATAGAACTCTATAATGCAACCCGAGGGGC 1750
```

```
OY 1681 GATACAAGCAGGATGTGTACTTCTCTCTAAGAAAATGATGAATACGTGGCCAGCTTGC 1740
    |||
Db 1751 GATACAAGCAGGATGTGTACTTCTCTCTAAGAAAATGATGAATACGTGGCCAGCTTGC 1810
OY 1741 ATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1800
    |||
Db 1811 ATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1870
OY 1801 GACTCAACAAAATGGGCCATTCAAACCTAATATTATACAGATAC 1844
    |||
Db 1871 GACTCAACAAAATGGGCCATTCAAACCTAATATTATACAGATAC 1914
```

```
RESULT 4
AAV73924
ID AAV73924 standard; DNA; 3616 BP.
XX
AC AAV73924;
XX
DT 04-MAR-1999 (first entry)
XX
DE Human SAHH DNA #1.
XX
KW S-adenosyl-5-homocysteine hydrolase; SAHH; human; drug screening;
KW treatment; infection; cancer; autoimmune disease; detection; diagnosis;
KW gene mapping; antisense; therapy; antagonist; immunoassay; ss.
XX
OS Homo sapiens.
XX
FH Key 56.1558
FT CDS /*tag- a
FT /product- "SAHH"
XX
XX US58540235A
XX
XX 29-DEC-1998.
XX
XX 17-JUL-1997; 97US-0896005.
XX
XX 17-JUL-1997; 97US-0896005.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Corley NC, Hillman JL, Lal P, Shah P;
XX
XX WPI; 1999-094906/08.
XX
XX P-PSDB; AAW90061.
XX
XX
XX Nucleic acid encoding human S-adenosyl-5-homocysteine hydrolase
XX for production of recombinant enzyme, useful for diagnosis,
XX treatment and prevention of cancers, infections and autoimmune
XX diseases
XX
XX Disclosure; Fig 1A-I; 40pp; English.
XX
XX
XX This sequence encodes a human S-adenosyl-5-homocysteine hydrolase (SAHH).
XX The SAHH protein can be used to generate specific antibodies and in drug
XX screening to identify specific binding agents. Antagonists of the
XX protein are used to treat or prevent a wide range of viral, bacterial,
XX fungal, parasitic, protozoal or helminthic infections, many cancers
XX (leukemia, lymphoma or solid tumors), and many autoimmune diseases
XX (e.g. acquired immune deficiency syndrome, allergy, asthma, diabetes
XX mellitus, multiple sclerosis etc). All these conditions may be treated by
XX expressing antisense sequences, triplex-forming agents or ribozymes
XX directed against the nucleic acid. The nucleic acid and its fragments can
XX be used as probes or primers for detecting and quantifying gene
XX expression, for diagnosis or monitoring of disease, to identify genetic
XX variations, mutations or polymorphisms, in gene mapping and as antisense
XX therapeutics. Antibodies are used directly as antagonists, indirectly to
XX deliver active agents to SAHH-expressing cells, to diagnose and monitor
XX diseases in standard immunoassays, in competitive drug screens and to
XX isolate the protein from natural sources.
```

xx Sequence 3616 BP; 1017 A; 782 C; 826 G; 991 T; 0 other;
SQ
Query Match 84.2%; Score 1551.8; DB 20; Length 3616;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1553; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 290 CGGGAGAGAGCTGAGCAGGCGCAAGAGATCGAGAGCGCGGAGAGTACTCTTCATGCG 349
DB 1 CAGGAGAGAGCTGAGCAGGCGCAAGAGATCGAGAGCGCGGAGAGTACTCTTCATGCG 60
QY 350 CACCGTCACCAAGCGCGCAAGAGCAAAATCCAGTTTGTGATGACATGCAGAGTTTCA 409
DB 61 CACCGTCACCAAGCGCGCAAGAGCAAAATCCAGTTTGTGATGACATGCAGAGTTTCA 120
QY 410 CAAATTCGCCCAACCAAACTGGCCCAAGATCTTGTCTCGCTGATCTCACAGTCTCCAC 469
DB 121 CAAATTCGCCCAACCAAACTGGCCCAAGATCTTGTCTCGCTGATCTCACAGTCTCCAC 180
QY 470 TGACAGCTACAGTTCAGCTGCATCTTACACAGATAGCTGTGATGAGGTTTCTCCCG 529
DB 181 TGACAGCTACAGTTCAGCTGCATCTTACACAGATAGCTGTGATGAGGTTTCTCCCG 240
QY 530 AGAGAGCAGCAGCAAACTCCAAAGGCGCAGCAGCAATTTCTGTGTGAAGAACATCAAGCA 589
DB 241 AGAGAGCAGCAGCAAACTCCAAAGGCGCAGCAGCAATTTCTGTGTGAAGAACATCAAGCA 300
QY 590 GGCAGATTTGGACGCGCGGAGATGAGATTGACAGAGCAAGATGTCTGTGATTTTC 649
DB 301 GGCAGATTTGGACGCGCGGAGATGAGATTGACAGAGCAAGATGTCTGTGATTTTC 360
QY 650 ACTCAGGAAACGTGCTCAGGCGGAGAGAGCCCTTGCTGTGCTAAATAGTGGGCTGTAC 709
DB 361 ACTCAGGAAACGTGCTCAGGCGGAGAGAGCCCTTGCTGTGCTAAATAGTGGGCTGTAC 420
QY 710 ACACATCACAGCCAGACAGCGGTGTGATTGAGACACTGTGCCCTGGGGCTCAGTG 769
DB 421 ACACATCACAGCCAGACAGCGGTGTGATTGAGACACTGTGCCCTGGGGCTCAGTG 480
QY 770 CCGCTGGTCTGCTGTGTAACATCTACTCAACTCAGATGAGTACGTGCAGCACTGGCTGA 829
DB 481 CCGCTGGTCTGCTGTGTAACATCTACTCAACTCAGATGAGTACGTGCAGCACTGGCTGA 540
QY 830 GCGTGAGTGTGAGTGTTCGCTTGGAAGGGGAGTCAAGATGACTTCTGTGTGTAT 889
DB 541 GCGTGAGTGTGAGTGTTCGCTTGGAAGGGGAGTCAAGATGACTTCTGTGTGTAT 600
QY 890 TGACCGCTGTGTGAACATGATGGCTGGCAGGCGCAACATGATCTGTGATGGGGAGA 949
DB 601 TGACCGCTGTGTGAACATGATGGCTGGCAGGCGCAACATGATCTGTGATGGGGAGA 660
QY 950 CTTAACCCACTGGTTTATAGAAGTATCCAAACGTGTTAAGAAGATCCGAGGATGT 1009
DB 661 CTTAACCCACTGGTTTATAGAAGTATCCAAACGTGTTAAGAAGATCCGAGGATGT 720
QY 1010 GGAAGAGAGCGTGAAGTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTTG 1069
DB 721 GGAAGAGAGCGTGAAGTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTTG 780
QY 1070 TGTTCGGGCGCATGAACGTCAATGATCTGTATCCAAACAGAGTTGATACTTGTACTG 1129
DB 781 TGTTCGGGCGCATGAACGTCAATGATCTGTATCCAAACAGAGTTGATACTTGTACTG 840
QY 1130 CTGCCGAGAAATCCATTTGGATGGCTGAAGAGAGACACAGATGTGATGTTGGTGA 1189
DB 841 CTGCCGAGAAATCCATTTGGATGGCTGAAGAGAGACACAGATGTGATGTTGGTGA 900
QY 1190 ACAAGT 1249
DB 901 ACAAGT 960
QY 1250 TCTTGAGCAATGTCTACATTACCGAAATGACCCCATCTGTGCTGCAGGCGCTGAT 1309
DB 1250 TCTTGAGCAATGTCTACATTACCGAAATGACCCCATCTGTGCTGCAGGCGCTGAT 1309

DB 961 TCTTGAGCAATGTCTACATTACCGAAATGACCCCATCTGTGCTGTGCAGGCGCTGAT 1020
QY 1310 GGATGGGTTACAGGTTGTTAAAGCTAATGAAGTATCTCCGGCAAGTCGATGCTAATAAC 1369
DB 1021 GGATGGGTTACAGGTTGTTAAAGCTAATGAAGTATCTCCGGCAAGTCGATGCTAATAAC 1080
QY 1370 TTGCACAGGAATAAGATGTAGTACACGGGAGCACTTGGATCGCATGAAAAACAGTTG 1429
DB 1081 TTGCACAGGAATAAGATGTAGTACACGGGAGCACTTGGATCGCATGAAAAACAGTTG 1140
QY 1430 TATCGTATGCAATATGGGCGCACTCCAAACAGAAATCGATGTGACAGCCTCCGCACTCC 1489
DB 1141 TATCGTATGCAATATGGGCGCACTCCAAACAGAAATCGATGTGACAGCCTCCGCACTCC 1200
QY 1490 GGAGTGAAGTGGGAGCGAGTACGTTCTCAGGTGACCATGTGATCTGGCCAGATGGCAA 1549
DB 1201 GGAGTGAAGTGGGAGCGAGTACGTTCTCAGGTGACCATGTGATCTGGCCAGATGGCAA 1260
QY 1550 ACAGTGTCT 1609
DB 1261 ACAGTGTCT 1320
QY 1610 CTTGTCT 1669
DB 1321 CTTGTCT 1380
QY 1670 ACCCGAGGCGGATACAGCAGAGTGTGATCTGCTTCTCTAAGAAATGATGAATACGT 1729
DB 1381 ACCCGAGGCGGATACAGCAGAGTGTGATCTGCTTCTCTAAGAAATGATGAATACGT 1440
QY 1730 TGCCAGCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCA 1789
DB 1441 TGCCAGCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCA 1500
QY 1790 AAAATATCTGGGACTCAACAAAAATGGGCCATTCAAACTTAATTATACAGATAC 1844
DB 1501 AAAATATCTGGGACTCAACAAAAATGGGCCATTCAAACTTAATTATACAGATAC 1555
RESULT 5
AAS76216/c
ID AAS76216 standard; cDNA; 5030 BP.
XX AC AAS76216;
XX AC
XX DT 13-FEB-2002 (first entry)
XX DE
XX DNA encoding novel human diagnostic protein #12020.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI: 2001-639362/73.
XX DR P-PSDB: ABG12029.
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions
XX
PS Claim 1; SEQ ID NO 40834; 21bp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (AB157737-AB172072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1776 BP; 437 A; 432 C; 479 G; 428 T; 0 other;

Query Match 35.2%; Score 649.2; DB 23; Length 1776;
Best Local Similarity 67.0%; Pred. No. 3.3e-163;
Matches 937; Conservative 0; Mismatches 458; Indels 3; Gaps 1;

QY 445 CTCGCTGATCTACAGTCTCCACTGACAGCTACAGTTCAGTGCATCTACACAGATA 504
DB 275 CTCGGCGGTTTCTGATGTCACACCTGCTCATCTCAGTTCCACCTGTTCCACCGAGACT 334
QY 505 GCTCTGATGATGAGTTCTCCCGGAGAGAGACGACCAACCACTCCAGGGCAGAGACA 564
DB 335 CCGACGAGGAGAGAGCTCTCCCGGAGAGACCAACCACTCCCGCGGAGAGACTG 394
QY 565 ATTCTGTGTGAGAGACATCAAGCAGGAGATTTGACGCGCGGAGATTGAGATTGACG 624
DB 395 ACTTTGCGTGAAGAGACTCTGAGAGAGTGCATTTGAGAGCGGAGATCGAGATCGCG 454
QY 625 AGCAAGACATGTCTGCTGATTTCACTCAGGAAACGCTCAGGGGAGAGCCCTTGG 684
DB 455 AGTCGAGATGCCGGGATCATGACTCTGAGGAAAGAGCGAAGATGAGAAAGCCCTAA 514
QY 685 CTGCTGCTAAATAGTGGGCTGTACACACATCACAGCCAGACAGCGGTGTGATTGAGA 744
DB 515 AGGGTGCATATCTGCGATGACACCCAGCTCAATGCTCAGTCGGCAGTGTGATCGAGA 574
QY 745 CACTCTGTCCCTGGGGCTCAAGTCCGCTGCTGTTTAACATCTACTCACTCAGA 804
DB 575 CCCTCGTCCAACTGGGGCCACAGTTCGCTGGGCTGCTGCAACATTTATTCACACAAA 634
QY 805 ATGAGTAGCTGACAGACTGGCTGAGGCTGAGTTCAGTTCGCTTGAAGGCGAGT 864
DB 635 ACGCAGTTGCCCGCTCTGGCAGAGCGGGAATCCGATCTTCGCTGGCGGAGAGA 694
QY 865 CAGAAGATGACTTCTGCTGCTGTATTGACCGCTGTGTAACATGATGGGTGGAGCCA 924
DB 695 CGGAGGAGAGAGTCTGCTGCTGCTTGGACAGGGCCATCTACTCCGACGGCTGGACCCA 754
QY 925 ACATGATCTGATGATGGGGGAGACTTAACCCACTGGGTTTATAAGAGTATCCAAAG 984
DB 755 ACCTAATCTGACGAGCGGGCGGATGCCACGCACTTATGCTCAAGAGTACCCGACT 814
QY 985 TGTTTAAGAGATCCGAGCATTTGTGAGAGAGAGCGTGTGTTACAGAGCTGTATC 1044
DB 815 ACTTCAAGGCATTCGGGCACTGCTGAGGAAAGTGAACGGGGTGCACCGCTGTACA 874
QY 1045 AGCTCTCCAAGCTGGGAGCTCTGTGTTCCGGCCATGAACGTCATGATCTGTACCA 1104
DB 875 TGCTGTCAAAGGGCGGAAACTTACTGTTCCGGCCATCAACGTTAACGACTCAGACCA 934
QY 1105 AACAGAGTTTGATTAATCTGACTGCTCCGAGAAATCCATTTTGATGGCCTGAGAGA 1164
DB 935 AGAACAAGTTTGATTAATCTGACTGCTCCGAGAAATCCATTTTGATGGCCTGAGAGA 994
QY 1165 CCACAGATGTGATGTTGCTGGGAAACAAGTGTGTGTGCTATGCTGAGGTAGGCA 1224

DB 995 CCACGGATATATGTTTGGCGGAAAGCAGGTGCTGATCTGTGGGTACGGTGTATGGGAA 1054
QY 1225 AGGCTGCTGTGCTGCTCTCAAGCTCTTGAGAGCAATGCTCTACATTACCGAAATGACC 1284
DB 1055 AGGGCTGTGCCAGTCTCCGGAAGGGCCAGAGATGATGTTTATGTTACGGAAGTGATC 1114
QY 1285 CCACTGTGCTCTGACAGGCTGATGATGGGTTTACAGGTTGTAAGCTTAATGAGTCA 1344
DB 1115 CCATATGTGCTCTACAAAGCTGCCATGATGATGATTCGGGTGTACGCTCAACGAGTCA 1174
QY 1345 TCCGCAAGTGCATGCTGTAATACTTGACACAGAAATAAGATGTAGTACAGCGGAGC 1404
DB 1175 TCAGGACGGTGTGATGCTGTGTTACGGCACTGGAACAAATAATGTTATTACAGGAGTC 1234
QY 1405 ACTTGATCGCATGAATAAGTGTATGCTATGCAATATGGGCCACTCCACACAGAAA 1464
DB 1235 ACATGAATCGCATGAAGATGTTGTATCTCTGCAATATGGGACATTCCTGCTCGAGA 1294
QY 1465 TCATGTGACACAGCTCCGCACTCCGAGCTGACGTTGGAGCGAGTCTCTCAGTGG 1524
DB 1295 TTGATGTGAATGCTTGCATACCCCGAGCTAACGTTGGAGCGTGTCCGTTCTCAGTGG 1354
QY 1525 ACCATGTCTATGCGCCAGATGGCAACAGAGTGTCTCTCTGCGAGAGGCTGCTACTCA 1584
DB 1355 ATCAGATCAGGTGGCGCGAGCGGAGAGATGATCATTTGCTCGCGAGGAGAGACTGTA 1414
QY 1585 ATTGAGCTGCTCCACAGTTCACCTTTGTTCTGCTATCAGACACACACAGCTT 1644
DB 1415 ATTTGCTGCTGTCACCACTTCGTCCTTTGCTGATCGGTGGCTCATCCACAGGCTT 1474
QY 1645 TGGCACTGATAGACTCTAATATGCAACCCGAGGGGCAATACAAGAGATGTACTTGC 1704
DB 1475 TGGCCCTGATGTAACCTCTCTCAGCGCC--AGGAAGATATAAGTCGATGTCTACTGC 1531
QY 1705 TTCTTAAGAAATGATGATATACGTTGCCAGCTTGCATCTGCCATTTGATGCCACC 1764
DB 1532 TGCCAAAGAAATGATGATATACGTTGCCAGTTGTCATCTGCCACCTTCGATGCTATC 1591
QY 1765 TTACAGAGCTGACAGATGACCAAGCAAAATATCTGGAGCTCAACAAATGGCCATTCA 1824
DB 1592 TCACGAGCTCACGAGATGACAGTCCAAAGTTTATGGCTTAAACAAGCGCGGCTTTA 1651
QY 1825 AACCTAATTATTACAGAT 1842
DB 1652 AAGCAATTAATAAGAT 1669

RESULT 9
AB103692/c
ID ABL03692 standard; cDNA; 4406 BP.
XX
AC ABL03692;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 5558.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.

CC ABL87934, (III) encoding (II) having a sequence (S2), a T cell
CC population of (II), or antigen presenting cells that express (II).
CC (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to
CC (S1) can be used for detecting ovarian cancer in a patient's biological
CC sample preferably serum or ovarian tissue. The method comprises
CC contacting a biological sample from a patient with (IV), detecting the
CC amount of polynucleotide hybridising to (IV) and comparing the amount to
CC a predetermined cutoff value and thereby detecting ovarian cancer in the
CC patient, where the amount of polynucleotide hybridising to (IV) is
CC detected preferably by polymerase chain reaction (PCR). (I) comprising
CC (III) and/or (II) is useful for stimulating and/or expanding T cells
CC specific for an ovarian tumour protein comprising contacting T cells
CC with (III) or (II). (III) is useful in design and preparation of
CC ribozyme molecules for inhibiting expression of the tumour polypeptides
CC and proteins in tumour cells; and to isolate a full length gene from a
CC suitable library e.g., a tumour cDNA library using well known
CC techniques.

XX
XX
SQ Sequence 553 BP; 159 A; 136 C; 128 G; 127 T; 3 other;

Query Match 25.0%; Score 461.4; DB 24; Length 553;
Best Local Similarity 97.3%; Pred. No. 3.3e-113;
Matches 510; Conservative 0; Mismatches 9; Indels 5; Gaps 4;

OY 1322 GGTGTAAGCTAATGAAGTCATCCGGCAAGTCATGCTGTAATACTTGACAGGAAA 1381
DB 1 GGTGTAAGCTAATGAAGTCATCCGGCAAGTCATGCTGTAATACTTGACAGGAAA 60
OY 1382 TAAGATGTAGTGACACGGGAGCAGCTGGATCGATGAAAACAGTTGATCGTATGCAA 1441
DB 61 TAAGATGTAGTGACACGGGAGCAGCTGGATCGATGAAAACAGTTGATCGTATGCAA 120
OY 1442 TATGGGCCACTCCACACAGAAATGATGTGACAGCCTCCGACCTCGAGCTGACGTG 1501
DB 121 TATGGGCCACTCCACACAGAAATGATGTGACAGCCTCCGACCTCGAGCTGACGTG 180
OY 1502 GGAGCGAGTACGTTCTCAGTGGACCATGTCATCTGGCCAGATGGCAACGAGTTGTCT 1561
DB 181 GGAGCGAGTACGTTCTCAGTGGACCATGTCATCTGGCCAGAT-GNAAACGAGTTGTCT 239
OY 1562 CCTGCAGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTCCCACTTTTGTCTGTC 1621
DB 240 CCTGCAGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTCCCACTTTTGTCTGTC 299
OY 1622 CATCACAGCCACACACAGGCTTTGGCAGTATAGAACTCTATAATGACCCGAGGGCG 1681
DB 300 CATCACAGCCACACACAGGCTTTGGCAGTATAGAACTCTATAATGACCCGAGGGCG 359
OY 1682 ATACAAGCAGAGTGTCTACTTGTCTTCTTAAGAAAA-TGGATGAATACGTTGCCAGCTTGC 1740
DB 360 ATACAAGCAGAGTGTCTACTTGTCTTCTTAAGAAAAATGGATGAATACGTTGCCAGCTTGC 419
OY 1741 ATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1800
DB 420 ATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGC-AAATATCTGG 478
OY 1801 GACTCAACAAAAATGGCCATTCAAACCTTAATTTATACAGATAC 1844
DB 479 GACTCAACAAAA--TGGGCCATTCAAACCTTAATTTATACAGATAC 520

RESULT 11
ABV44141/C
ID ABV44141 standard; cDNA; 504 BP.
XX AC ABV44141;
XX 16-SEP-2002 (first entry)
DE Human prostate expression marker cDNA 44132.
XX Human prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.

XX
OS Homo sapiens.
XX WO200160860-A2.
XX 23-AUG-2001.
XX 20-FEB-2001; 2001WO-US05171.
XX 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Schlegel R, Endege WO, Monahan JE;

WPI; 2001-662795/76.

Novel isolated nucleic acid molecule associated with cancerous state of
prostate cells and correlating with presence of prostate cancer, useful
for detecting presence of prostate cancer, stage of prostate cancer

Claim 1; Page 8770; 11750pp; English.

XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX
SQ Sequence 504 BP; 117 A; 150 C; 116 G; 121 T; 0 other;

Query Match 22.9%; Score 422; DB 23; Length 504;
Best Local Similarity 100.0%; Pred. No. 1.1e-102;
Matches 422; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 706 GTACACACATCACAGCCCAAGACAGCGGTGATGAGACACTCTGCCCCGCGGCTC 765
DB 502 GTACACACATCACAGCCCAAGACAGCGGTGATGAGACACTCTGCCCCGCGGCTC 443
OY 766 AGTCCCGCTGCTCTGCTTGTACATCTACTCACTCAGAAATGAAGTAGCTGACGACTGG 825
DB 442 AGTCCCGCTGCTCTGCTTGTACATCTACTCACTCAGAAATGAAGTAGCTGACGACTGG 383
OY 826 CTGAGGCTGAGTGTGAGTGTGCTTGGAGGGCGAGTCAGAGATGACTTCTGTGTGT 885
DB 382 CTGAGGCTGAGTGTGAGTGTGCTTGGAGGGCGAGTCAGAGATGACTTCTGTGTGT 323
OY 886 GTATTGACCGCTGTGTGAACATGATGGGTGCGAGGCCAACATGATCCTGGATGATGGG 945
DB 322 GTATTGACCGCTGTGTGAACATGATGGGTGCGAGGCCAACATGATCCTGGATGATGGG 263
OY 946 GAGACTTAACCCACTGGGTTTATAAGAGATATCCAAACGTGTTTAAGAGATCCGAGGCA 1005
DB 262 GAGACTTAACCCACTGGGTTTATAAGAGATATCCAAACGTGTTTAAGAGATCCGAGGCA 203
OY 1006 TTGTGAAGAGAGCGTGAAGTGTGTTACAGAGCTGTATCAGCTCTCCAAAGCTGGGAAGC 1065
DB 202 TTGTGAAGAGAGCGTGAAGTGTGTTACAGAGCTGTATCAGCTCTCCAAAGCTGGGAAGC 143


```

Db      4026 AGCGTGCCGTTCTCAAGTGGATCATCATCAGTGGCCGCGACGACGAGATGATCATTTTGC 4085
QY      1564 TGGCAGAGGGTGTCTACTCAATTTGAGCTGCTCCACAGTCCACCTTTGTCTGTCCA 1623
        ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db      4086 TCGCCGAGGGAAGACTGTGAATTTGTCTGTCCACCACTTTCCTTTGTGCGATCCG 4145
QY      1624 TCACAGCCACAACAGAGCCTTTGGCACTGATAGACTCTATAATGCACCCGAGGGCGAT 1683
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      4146 TGGCCTCATCCACCCAGCCTTTGGCCCTGATTGAACCTTCTCAGCGCC--AGGAAGAT 4202
QY      1684 ACAAGCAGGATGTGTACTTGTCTCTCCTAAGAAATGATGATACGTTG 1731
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      4203 ATAGTCGGATGTCTACCTGCTGCCAAGAAATGGGTAGCTTCTTAG 4250

RESULT 13
AAZ80766/c
ID      AAZ80766 standard; cDNA; 636 BP.
XX
AC      AAZ80766;
XX
DT      07-APR-2000 (first entry)
XX
DE      Human colon cancer cell line SW480 cDNA clone SEQ ID NO:850.
XX
KW      Human; gene expression product; diagnosis; tumour; colon cancer;
KW      colorectal adenocarcinoma; cell line SW480; cell proliferation;
KW      cytosstatic; sarcoma; breast cancer; neoplasia; dysplasia;
KW      hyperplasia; ds.
XX
OS      Homo sapiens.
XX
PN      WO9964576-A2.
XX
PD      16-DEC-1999.
XX
PE      09-JUN-1999; 99WO-IB01062.
XX
PR      10-JUN-1998; 98US-0088801.
XX
PA      (FARB ) BAYER CORP.
XX
PI      Endege WO, Steinmann KE, Astle JH, Burgess CC, Bushnell SE;
PI      Carroll E, Catino TJ, Dertl A, Ford DM, Lewis ME, Monahan JE;
PI      Schlegel R;
XX
DR      WPI; 2000-087220/07.
XX
PT      Novel nucleic acids, used to develop products for the diagnosis and
PT      treatment of disorders involving unwanted cell proliferation,
PT      particularly cancers, especially colon cancer
XX
PS      Claim 15; Page 469; 469pp; English.
XX
AAZ79917 to AAZ80766 represent double stranded cDNA clones isolated from
CC      the human colorectal adenocarcinoma (colon cancer) cell line SW480. The
CC      cDNA clones can be used to generate antisense oligonucleotides which
CC      can be used for antisense therapy. Methods and products from the present
CC      invention can be used for identifying and/or classifying cancerous cells
CC      present in a human tumour, particularly in solid tumours, e.g.
CC      carcinomas and sarcomas, e.g. breast or colon cancers. The cDNA clones
CC      can be used for developing agents for the diagnosis and treatment of
CC      disorders involving unwanted cell proliferation, such as neoplasia,
CC      dysplasia or hyperplasia.
XX
SQ      Sequence 636 BP; 156 A; 166 C; 137 G; 147 T; 30 other;

Query Match          22.0%; Score 406.4; DB 21; Length 636;
Best Local Similarity 99.5%; Pred. No. 1.8e-98;
Matches 418; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

707 TACACACATCACAGCCACAGACAGC-GGTGTGATTGAGACACTCTGTGCCCTGGGGGCTC 765
|||||

```

Db	420	TACACACATCACAGCCACAGACAGCGGCTGTGATTGAGACACTCTGTGCTCTGCGGGCTC	361
QY	766	AGTCCCGCTGCTCTGCTTGTAAACATCTACTCAACTCAGAATGAAGTAGCTGCAGCACTGG	825
Db	360	AGTCCCGCTGCTCTGCTTGTAAACATCTACTCAACTCAGAATGAAGTAGCTGCAGCACTGG	301
QY	826	CTGAGCGCTGAGTTGCAGTGTTCGCTTGGAAAGGCGAGTCAGAAAGTACTTCTGTGTGT	885
Db	300	CTGAGCGCTGAGTTGCAGTGTTCGCTTGGAAAGGCGAGTCAGAAAGTACTTCTGTGTGT	241
QY	886	GTAATTGACCGCTGTGTGAACATGATGGGTGGCAGGCCAACATGATCCTGGATGATGGG	945
Db	240	GTAATTGACCGCTGTGTGAACATGATGGGTGGCAGGCCAACATGATCCTGGATGATGGG	181
QY	946	GAGACTTAACCCACTGGGTTTATAGAAGTATCCAAACGTGTTAAGAAGATCCGAGCA	1005
Db	180	GAGACTTAACCCACTGGGTTTATAGAAGTATCCAAACGTGTTAAGAAGATCCGAGCA	121
QY	1006	TTGTGGAAGAGAGCGTGACTGTGTTTACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGC	1065
Db	120	TTGTGGAAGAGAGCGTGACTGTGTTTACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGC	61
QY	1066	TTCTGTGTCCCGCATGAACGTCAATGATTTCTGTTACCAACAGAAGTTGATAACTTGT	1125
Db	60	TTCTGTGTCCCGCATGAACGTCAATGATTTCTGTTACCAACAGAAGTTGATAACTTGT	1

CC	PT	DT	AC	ID	RESULT 14
CC	PT	DT	AC	ID	AAC77823
CC	PT	DT	AC	ID	AAC77823 standard; cDNA; 2200 BP.
CC	PT	DT	AC	ID	AAC77823;
CC	PT	DT	AC	ID	08-FEB-2001 (first entry)
CC	PT	DT	AC	ID	Human cancer associated gene sequence SEQ ID NO:217.
CC	PT	DT	AC	ID	Human; cancer associated gene; cancer antigen; detection; cancer;
CC	PT	DT	AC	ID	diagnosis; cytostatic; proliferative; vulnerrary; immunomodulator;
CC	PT	DT	AC	ID	antidiabetic; antilasthmatic; antirheumatic; antiarthritic; antiviral;
CC	PT	DT	AC	ID	antiinflammatory; antithyroid; antiallergic; antibacterial; cardiac;
CC	PT	DT	AC	ID	dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
CC	PT	DT	AC	ID	vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation
CC	PT	DT	AC	ID	immune disorder; haematopoietic cell disorder; autoimmune disorder;
CC	PT	DT	AC	ID	allergic reaction; graft versus host disease; organ rejection;
CC	PT	DT	AC	ID	haemostatic; thrombolytic; cardiovascular disorder; infection;
CC	PT	DT	AC	ID	neurological disease; drug screening; ss.
CC	PT	DT	AC	ID	Homo sapiens.
CC	PT	DT	AC	ID	WO200055350-A1.
CC	PT	DT	AC	ID	21-SEP-2000.
CC	PT	DT	AC	ID	08-MAR-2000; 2000WO-US05882.
CC	PT	DT	AC	ID	12-MAR-1999; 99US-0124270.
CC	PT	DT	AC	ID	(HUMA-) HUMAN GENOME SCI INC.
CC	PT	DT	AC	ID	Rosen CA, Ruben SM;
CC	PT	DT	AC	ID	WPI; 2000-587533/55.
CC	PT	DT	AC	ID	P-PSDB; AAB43614.
CC	PT	DT	AC	ID	Novel isolated nucleic acids comprising sequences encoding peptides
CC	PT	DT	AC	ID	useful for treating or diagnosing e.g. cancer -
CC	PT	DT	AC	ID	Claim 1; Page 787-788; 2352pp; English.
CC	PT	DT	AC	ID	AAC77607 to AAC78448 encode the human cancer associated proteins given
CC	PT	DT	AC	ID	in AAB43398 to AAB44239. The proteins can have activities based on the
CC	PT	DT	AC	ID	tissues and cells the genes are expressed in. Example of activities

CC include: cytostatic; proliferative; vulnerary; immunomodulator;
 CC antidiabetic; antiasmatic; antirheumatic; antiarthritic;
 CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
 CC nootropic; vasotropic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention.

XX Sequence 2200 BP; 546 A; 578 C; 607 G; 466 T; 3 other;

Query Match 21.8%; Score 402.4; DB 21; Length 2200;
 Best Local Similarity 58.4%; Pred. No. 3.9e-97;
 Matches 741; Conservative 0; Mismatches 521; Indels 6; Gaps 2;

QY 580 ACATCAGCAGGAGGAGATTTGGACGCCGAGATGAGATTGCAGAGCAAGACATGCTG 639
 DB 108 ACATCGCGCTGGCTGCTCGGGAGCGAGGCCCTGACATTTGCTGAGAACAGATGCCGG 167
 QY 640 CTCTGATTTCACTCAGGAACGTGCTCAGGGGGAGAGAGCCCTGGCTGGCTAAATAG 699
 DB 168 GCGTATGCGTATGCGGGAGCGGTACTCGCCCTCAAGCCACTGAAGGCGCCCGCATCG 227
 QY 700 TGGGCTGTACACATCACAGCCAGACGCGTGTGATGAGACACTCTGTCCTG 759
 DB 228 CTGGCTGCTGACATGACCGTGAAGGCGCTCTCATTTGAGACCCCTGTCACCTGG 287
 QY 760 GGGCTCAGTGCCTGCTGCTGTACATCTACTCACTCAGAAATGAAGTAGTCAG 819
 DB 288 GTGCTGAGGTGCAAGTGTCCAGCTGCAACATCTTCTCCACCCAGGACCATGCGGCGCTG 347
 QY 820 CACTGCTGAGGCTGGAGTGTGAGTGTGCTTGAAGGGCGAGTCAAGATGACTTCT 879
 DB 348 CCATGCGCAAGCGTGGCATTCGGGTATGCTTGAAGGGCGAAGCGAGAGAGTACC 407
 QY 880 GGTGCTATTTGACCGCTGTGTAACATGATGGGTGGCAGGCCAATGATCCTGATG 939
 DB 408 TGTGTGATTTGAGCAGACACCGCTGTACTCAAGGAGCGGCCCTCAACATGATCTGAGC 467
 QY 940 ATGGGAGAGACTTAACCACTGGGTTTATAAGAGTATCCAAACGTGTTTAAAGATCC 999
 DB 468 ACGGGGCGAGCTCACCAACCTCATCCACCAAGTACCCGAGCTTCTGCGCAGCATCC 527
 QY 1000 GAGGCTTGTGAAGAGAGCGTGTGCTGTCAAGGCGTGTATCACTCTCCAAAGCTG 1059
 DB 528 GAGGCACTCTGAGAGAGACCAAGAGTGGGTCCACAACCTTACAAGATGATGCCAATG 587
 QY 1060 GGAAGCTCTGTGTCGGCCATGAAGTCAATGATTTCTGTACCAACAGAAAGTTGATA 1119
 DB 588 GGATCCTCAAGGTGCTGCTCATCAATGATGATGACTCCGTACCAAGAGCAAGTTGACA 647
 QY 1120 ACTGTACTGCTGCCGAGATCCATTTGGATGGCTGGAAGAGGACACAGATGTGATGT 1179
 DB 648 ACCTTATGCTGCTGCCGAGAGTCCCTCATAGATGGCATCAAGCGGGCCACAGATGTGATGA 707
 QY 1180 TTGGTGGAAACAAGTGTGTGTGCTATGTTGAGTGAAGTGAAGGAGGCTGTGCTG 1239
 DB 708 TTGCCGGCAAGGTAGCGGTGTAGCAGGCTATGTGTGTTGGCAAGGCTGTGCCAGG 767
 QY 1240 CTCTCAAGCTCTTGAGCAATTTGTCAATTAACGAATGAGCCCACTGTGCTCTGC 1299
 DB 768 CCTGCGGGGTTCGAGACCGCGCTCATCATCACGAGATTGACCCCATCAACGCACTGC 827

QY 1300 AGCGTCATGATGGGTTACGGTGTGTAAGCTTAATGAAGTCATCCGCGCAAGTCGATG 1359
 DB 828 AGGCTGCCATGAGGGGCTATGAGGTGACCAACATGATGAGCGCTGCAGAGGCAACA 887
 QY 1360 TCGTAATACCTGACAGGAATGAATGATGACACGGAGCAGTGGATGCGATGA 1419
 DB 888 TCTTGTACACACACAGCGCTGATGTATGACATCATCCTTGGCCGCGACTGTGAGCATGA 947
 QY 1420 AAACAGTTGTATGATGCAATATGGCCCACTCCAAACAGAAATGATGTGACAGCC 1479
 DB 948 AGGATGATGCCATTTGTGTATGATGACATTTGACGTGAGATGATGATGATGATGATG 1007
 QY 1480 TCCGCACTCCGAGCTGACGTGGAGCGAGTACGTCTCAGGTGACCATGTCATCTGGC 1539
 DB 1008 TCAACGAGAACCGCGTGAAGAGTGAACATCAAGCCGAGGTGACCGGTATCGGTTGA 1067
 QY 1540 CAGATGGCAACAGTGTCTCTCTGCGCAGAGGCTGCTACTCAATTTGAGCTGCTCA 1599
 DB 1068 AGAATGGCGCGCATCATCTCTGCTGCGCGAGGGTCCGCTGCTCAACCTGGGTTGTCCA 1127
 QY 1600 CAG--TTCCACCTTGTCTGTCTCATCACAGCCACACACAGAGCTTGGCACTGATAG 1656
 DB 1128 TGGCCACCCAGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1187
 QY 1657 AACTCTAATGACACCCGAGGGCGATACAGAGATGTGACTTGTCTTCTTAAGAAA 1716
 DB 1188 AGCTGTGACCCATCCAGACAAG--TACCCGTTGGGTTTCAATTTCTGCGCAAGAAGC 1244
 QY 1717 TGGATGATACGTTCGCCAGCTTGCATCTGCCATCTTGTATGCCCCACCTTACAGAGCTGA 1776
 DB 1245 TGGATGAGGCAAGTGTGTAAGCCACCTGGGCAAGCTGATGTGAAGTGAACCAAGCTAA 1304
 QY 1777 CAGATGACCAAGCAAAATATCTGGGACTCAACAAAATGGCCATTCAAACCTTAATATT 1836
 DB 1305 CTGAGAAGCAAGCCAGTACCTGGGATGTCTGTGATGAGGCCCTTCAAGCGGATCACT 1364
 QY 1837 ACAGATAC 1844
 DB 1365 ACCGCTAC 1372

RESULT 15
 ABV23126
 ID ABV23126 standard; cDNA; 2658 BP.
 XX AC ABV23126;
 XX AC 16-SEP-2002 (first entry)
 DT 16-SEP-2002 (first entry)
 XX DE Human prostate expression marker cDNA 23117.
 XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX OS Homo sapiens.
 XX PN WO200160860-A2.
 XX PD 23-AUG-2001.
 XX PF 20-FEB-2001; 2001WO-US05171.
 XX PR 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX PI Schlegel R, Endege WO, Monahan JE;


```
Db 300 CTGAGGCTGAGTGTGACGTGCTGCTTGAAGGGGAGTCAAGATGACTTCTGTGCT 241
QY 886 GTATTGACCGCTGTGTGACATGATGGGTGGCAGGCCAACATGATCCTGATGATGGG 945
Db 240 GTATTGACCGCTGTGTGACATGATGGGTGGCAGGCCAACATGATCCTGATGATGGG 181
QY 946 GAGACTTAACCCACTGGGTTTATAAGAGTATCCAAACGTGTTTAAAGATCCGAGCA 1005
Db 180 GAGACTTAACCCACTGGGTTTATAAGAGTATCCAAACGTGTTTAAAGATCCGAGCA 121
QY 1006 TTGTGAAGAGAGCGTGTGCTGTCTACAGGCTGTATCAGCTCTCCAAAGCTGGGAGC 1065
Db 120 TTGTGAAGAGAGCGTGTGCTGTCTACAGGCTGTATCAGCTCTCCAAAGCTGGGAGC 61
QY 1066 TCTGTGTCCGGCCATGAGCTCAATGATCTGTATACCAACAGAGTTGATTAATCTGT 1125
Db 60 TCTGTGTCCGGCCATGAGCTCAATGATCTGTATACCAACAGAGTTGATTAATCTGT 1
```

RESULT 2

```
US-09-318-448-26
; Sequence 26, Application US/09318448
; Patent No. 6210950
; GENERAL INFORMATION:
; APPLICANT: Johnson, William G.
; APPLICANT: Stenroos, Edward S.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; FILE REFERENCE: 601-1-057
; CURRENT APPLICATION NUMBER: US/09/318,448
; CURRENT FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 26
; LENGTH: 2211
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-318-448-26
```

```
Query Match 17.0%; Score 314.4; DB 4; Length 2211;
Best Local Similarity 61.9%; Pred. No. 1.2e-72;
Matches 498; Conservative 0; Mismatches 306; Indels 0; Gaps 0;
```

```
QY 580 ACATCAAGCAGCAGAAATTGGACGCCGAGATTGAGATTGACAGCAGACATGCTG 639
Db 79 ACATCGGCTGGCTGCTGGGAGCGGAGCGGCTTCCAAAGCCACTGAAGGCCCGCATCG 138
QY 640 CTCTGATTCACTCAGAAACGTGCTCAGGGGAGAGAGCCCTGCTGCTAAATAG 699
Db 139 GCCTGATGGGTATGCGGAGCGGAGCTGCGGCTTCCAAAGCCACTGAAGGCCCGCATCG 198
QY 700 TGGGCTGTACACATCAGACGCCAGACGCGGTGTTGATTGAGACACTCTGCCCCG 759
Db 199 CTGGCTGCTGACATGACCGGTGAGAGCGGCGGCTTCCAAAGCCACTGAAGGCCCGCATCG 258
QY 760 GGGCTAGTGGCGCTGCTGCTGTGTAACATCTACTCACTCAGAAAGTAGCTGACG 819
Db 259 GTGCTGAGGTGAGTGTGCTCAGGTGCAACATCTTCTCCAGCAGAACCATGCGGCGCTG 318
QY 820 CACTGCTGAGGCTGGAGTTGAGTTCGCTTGAAGGGGAGAGTCAAGATGACTTCT 879
Db 319 CCATGCGCAAGGCTGGCATTCGGGTATGCTTGAAGGGGAGAGCGGAGAGGAGTACC 378
QY 880 GGTGTGTATGACCGCTGTGTGACATGATGGGTGGCAGGCCAACATGATCCTGATG 939
Db 379 TGTGTGATGATGAGACAGACCGGTGTATCTCAAGAGCGGCGGCTTCAACATGATCTGAGC 438
QY 940 ATGGGAGACTTAACCCACTGGGTTTATAAGAGTATCCAAACGTGTTTAAAGATCC 999
Db 439 ACGGGGCGAAGCTCAGCAACCTGATCCACCAAGTACCGGAGCTTCTGCGAGGATCC 498
QY 1000 GAGCATTTGTGAAGAGAGCGGTGCTGTGCTCAGGCTGTATCAGCTCTCCAAAGCTG 1059
Db 60 TCTGTGTCCGGCCATGAGCTCAATGATCTGTATACCAACAGAGTTGATTAATCTGT 1
```

```
Db 499 GAGGCATCTCTGAGAGAGACCGACTGGGGTCCACAACTCTTACAAGATGATGCCAATG 558
QY 1060 GGAAGCTCTGTGTCCGGCCATGAACTCAATGATCTGTATCCAAACAGAGTTGATA 1119
Db 559 GGATCCTCAAGGTGCTGCTGCTCAATGATGATGATGATGATGATGATGATGATGAT 618
QY 1120 ACTGTACTGCTGCGGAGATTCATTTTGGATGCGCTGAGAGAGACACAGATGATGT 1179
Db 619 ACCTTATGCTGCGGAGATTCCTCATATGATGATGATGATGATGATGATGATGATGAT 678
QY 1180 TTGTGGAAACAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1239
Db 679 TTGCGGCAAGTATGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 738
QY 1240 CTCTCAAGCTCTGTGAGCAATGTCTACATTAACCAATGCAACCCATCTGTCTGTG 1299
Db 739 CCTGCGGGGTTTCGAGCGCGGCTCATCATCATCACCAGATTTGACCCCATCAACGCTGC 798
QY 1300 AGGCTGCATGATGGGTTCAGGGGTGTTAAGCTTAATGAATGATCATCCGCAAGTGTG 1359
Db 799 AGGCTGCATGAGGGGTATGAGGTGACCAACCATGATGATGAGGCTGTGAGAGGCAACA 858
QY 1360 TCGTATTAATCTGCACAGAAATA 1383
Db 859 TCTTGTACACACACAGGCTGTA 882
```

RESULT 3

```
US-09-347-878-2
; Sequence 2, Application US/09347878C
; Patent No. 6376210
; GENERAL INFORMATION:
; APPLICANT: Yuan, Chong
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES
; FILE REFERENCE: 25885-1651
; CURRENT APPLICATION NUMBER: US/09/347,878C
; CURRENT FILING DATE: 1999-07-06
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2211
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Human S-adenosylhomocysteine hydrolase cDNA
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: M61831/GenBank
US-09-347-878-2
```

```
Query Match 17.0%; Score 314.4; DB 4; Length 2211;
Best Local Similarity 61.9%; Pred. No. 1.2e-72;
Matches 498; Conservative 0; Mismatches 306; Indels 0; Gaps 0;
```

```
QY 580 ACATCAAGCAGCAGAAATTGGACGCCGAGATTGAGATTGACAGCAGACATGCTG 639
Db 79 ACATCGGCTGGCTGCTGGGAGCGGAGCGGCTTCCAAAGCCACTGAAGGCCCGCATCG 138
QY 640 CTCTGATTCACTCAGAAACGTGCTCAGGGGAGAGAGCCCTGCTGCTAAATAG 699
Db 139 GCCTGATGGGTATGCGGAGCGGAGCTGCGGCTTCCAAAGCCACTGAAGGCCCGCATCG 198
QY 700 TGGGCTGTACACATCAGACGCCAGACGCGGTGTTGATTGAGACACTCTGCCCCG 759
Db 199 CTGGCTGCTGACATGACCGGTGAGAGCGGCGGCTTCCAAAGCCACTGAAGGCCCGCATCG 258
QY 760 GGGCTAGTGGCGCTGCTGCTGTGTAACATCTACTCACTCAGAAAGTAGCTGACG 819
Db 259 GTGCTGAGTGTGAGTGTGCTCAGCTGCAACATCTTCTCCAGCAGAACCATGCGGCGCTG 318
QY 820 CACTGCTGAGGCTGAGTGTGAGTGTGCTTGAAGGGGAGAGTCAAGATGACTTCT 879
Db 319 CCATGCGCAAGGCTGGCATTCGGGTATGCTTGAAGGGGAGAGCGGAGAGGAGTACC 378
```

OY	880	GGTGGTATATGACCCCGCTGTGTGAACATGGATGGGTGGCAGGCCAACATGATCCGTGATG	939
Db	379	TGTGTGCATTGAGCAGACCCTGTACTTCAAAGGACGGCCCCCTCAACATGATTCGGACG	438
OY	940	ATGGGGGAGACTTAACCCACTGGGTTTATAAGAAGTATCCAACGCTGTTAAGAAGATCC	999
Db	439	ACGGGGGGCACCTCACCAACCTCATCCACACCAAGTACCCCGACGCTCTGTCCAGGCATCC	498
OY	1000	GAGGCATGTGGAAGAGACCGTGACTGGTGTTCACAGGCTGTATCAGCTCTCCAAAGCTG	1059
Db	499	GAGGCATCTCTGAGAGAGACCAAGCACTGGGGTCCACAACCTCTACAAGATGATGGCCAATG	558
OY	1060	GGAAGCTCTGTGTTCCGGCCATGAACGTCATGATTTCTGTACCAAACAGAAGTTGATA	1119
Db	559	GGATCCTCAAGGTGCTTCCCATCATGTCAATGACTCCGTCCACCAAGACAAGTTTGACA	618
OY	1120	ACTTGTACTGCTGCCGAGAATCCATTTTGGATGGCCTGAAGAGACACAGATGTGATGT	1179
Db	619	ACCTCTATGGCTGCCGGGAGTCCCTCATAGATGGCATCAAGCGGGCCACAGATGTGATGA	678
OY	1180	TTGCTGGGAAACAAGTGTGTGTGTGTGGCTATGTGTGAGGTAGGCCAAGGGCTGCTGTGCTG	1239
Db	679	TTGCCGGCAAGGTAGCGGTGGTAGCAGGCTATGTGTGTGTGGGCAAGGGCTGTGCCCAGG	738
OY	1240	CTCTCAAAGCTCTTGGAGCAATTGCTACATTTACCAATTCAGACCCCATCTGTGCTGTC	1299
Db	739	CCCTGCGGGGTTTCGAGCCCGCGTCATCATCACCGAGATGACCCCATCAACGCACATGCG	798
OY	1300	AGGCCTGCATGATGGGTTCAGGGTGTAAAGCTAAATGAGTCAATCCGGCAAGTGCATG	1359
Db	799	AGGCTGCCATGGAGGGCTATGAGGTGACCAACCATGGATGAGGCCCTGTCAAGAGGGCAACA	858
OY	1360	TCGTAATACTTGCACAGGAATA 1383	
Db	859	TCTTTGTCAACCAACACAGGCTGTA 882	

```

RESULT 4
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

```

Query Match ₁	Best Local Similarity	Matches	Score	DB 4	Length	Mismatches	Indels	Gaps
970	AGAAGTATCCAAACGCTGTTTAAGAGATCCGAGGCATTGTGGAAGAGCGGTGACTGGTG	12.88;	236.2;	1029	4411529;	0;	9;	3;
3629036	ACAAAGTGGAACCAAGATAGCCGAGTCGGTCAAGGGCGTCAACCGAGAGAACCAACACCGGCG	56.68;	Pred. No. 9.2e-51;	3628977		0;		
1030	TTACACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTGTTCGGCCATGAACGTCA			1089				
3628976	TGCTCGCGGCTCTACCAATTGCGCGCGCGGGGATCTGGCCCTTCCCGGATCAACGTCA			3628917				
1090	ATGATTCTGTATCCAAACAGAAGTTGATTAACCTGTACTGCTGCCAGAAATCCATTTTGG			1149				

Db	3628916	ACGACTCGGTGACCAAGTCCAAATTCGACACAAAGTACGGCACTCCGCACTCCCTGATCG	3628857
QY	1150	ATGGCCCTGAAGAGGAGCCACAGATGTGATGTTTGGTGGGAAACAAGTGTGTGTGGCT	1209
Db	3628856	ACGGCATCAACCGCGGCACCGACGGCTGATCGGCGGTAGAGAAGTCCCTCATCTGCGGCT	3628797
QY	1210	ATGGTGAGGTAGGCAAGGGCTGTGCTGTGCTCTCTCAAGCTCTTGAGCAATTGTCTACA	1269
Db	3628796	ACGGCGACGTCGGTAAAGGCTGTGTGCGGAGGCGATGAAGGCCAGGGAGCGCGGCTCTCCG	3628737
QY	1270	TTACCGAAATCGACCCCATCTGTGCTCTGACAGGCTGCAATGATGGGTTGAGGGTGTAA	1329
Db	3628736	TCACCGAGATCGACCCGATCAACGCGCTGACAGGCCATGATGGAAGGCTTCGACGTGTCA	3628677
QY	1330	AGCTAAATGAAGTCATCCGGCAAGTGCATGTCTGTAATAACTTGCAACAGAAATGAAGATG	1389
Db	3628676	CGGTGAGGAGGCCATCGGGGAGCGCCGACATCGTGTAAACGCGACCGCAACAAGACA	3628617
QY	1390	TAGTGACACAGGAGCACACTTGGATCGCATGAAAAAACAGTTGTATCGTATGCAATATGGCC	1449
Db	3628616	TCATCATGCTCGAGCACATTAAGCCGATGAAGAGCACCGCGATCTGGGAAATATCGGCC	3628557
QY	1450	ACTCCACACACAGAAATCGATGTGACACAGCCTCCGCACTCCGAGCTGACGTGGAGCGAG	1509
Db	3628556	ACTTCGACACACAGAGATCGACATGCGCGCGGCTGGAGCGCTCCGGGGCGACACGGGTCAACG	3628497
QY	1510	TACGTTCTCAGGTGGAGCCATGTATCTGGCCAGA--TGCCAAACGAGTGTCTCTCTGG	1566
Db	3628496	TCAGCCTCAGGTCGACCTGTGACCTTTGGCGACACGGGCGCTCGATCATCTGTCTGT	3628437
QY	1567	CAGAGGGTGTCTACTCAATTTGAGCTGTCTCCACAGTTCACAC--TTTGTCTGTCCA	1623
Db	3628436	CGAGGGGCGGCTGTGTAACCTGGGCAATGCCACCGGGGACACCCCTGTTGTGATGAGCA	3628377
QY	1624	TCACAGCCACACACAGGCTTTGGCACTGATAGAACTCTATAATGCACCGGAGGGCGAT	1683
Db	3628376	ACAGCTTCGCTAACCAAGACGATGCGCCAGATCGAGCTGTGACCAAGAAGACGACGAG--T	3628320
QY	1684	ACAAGCAGGATGTACTTGTCTTCCCTAAGAAATGATGAATACGTTGCCAGCTTGCAATC	1743
Db	3628319	ACGACACAGAGGTGTACCGGCTGGCCAAAGCACCTCGACGAGAGGTGGCTCGAATCCATG	3628260
QY	1744	TGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGAC	1803
Db	3628259	TCGAGGCCCTTGGCGGTACCTGACCAAGCTGACCAAGAGCAGGCCGAATACCTCGGGC	3628200
QY	1804	TCACAAAAATGGGCATTCAAACCTAATATTACAGATAC	1844
Db	3628199	TCGACGTCGAGGGTCCCTACAAAGCCGGAACCACTACCGGCTAC	3628159

RESULT 5
US-08-669-536-1
; Sequence 1, Application US/08669536
; Patent No. 5910444
; GENERAL INFORMATION:
; APPLICANT: MASUTA, CHIKARA
; APPLICANT: UEHARA, KYOKO
; APPLICANT: TANAKA, HIDEO
; APPLICANT: KUWATA, SHIGERU
; TITLE OF INVENTION: ORGANISMS IN WHICH THE EXPRESSION OF
; TITLE OF INVENTION: S-ADENOSYLHOMOCYSTEINE HYDROLASE GENE IS INHIBITED
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30.
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/669,536
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: MURPHY JR, GERALD M
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1254-128
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1812 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-08-669-536-1

Query Match 12.1%; Score 222.6; DB 2; Length 1812;
Best Local Similarity 56.4%; Pred. No. 1.1e-48;
Matches 499; Conservative 0; Mismatches 374; Indels 12; Gaps 4;

QY 972 AAGTATCCAAAGCTGTTAAGAAGATCCGAGCATGTGGAAGAGAGCGTGTGTT 1031
DB 634 AATATACCAAGATGAAGAAAGACTGTCGGTCTTCTGAGAACTACCACTGAGTT 693
QY 1032 CACAGCGTGTATGAGCTCTCCAAAGCTGGGAAGCTCTGTCTCCGCCAATGAGTCAAT 1091
DB 694 AAGAGGCTTATGAGTGAAGGCTAATGGAATGCTTCTCCCTGCTATTAATGTTAAT 753
QY 1092 GATTCTGTACCAACAGAAAGTTGATTAATCTGTACTGCTGCCGAGAATCCATTTTGAT 1151
DB 754 GATTCTGTACCAACAGAAAGTTGATTAATCTGTACTGCTGCCGAGAATCCATTTTGAT 813
QY 1152 GGCCTGAGAGAGACCAAGATGTGATGTTGGTGGGAACAAGTGGTGTGTTGCTAT 1211
DB 814 GGTCTCATGAGGGCTACTGATGTTATGATGTCGGGAAGGTTGCCCTGTTGCTGTTAT 873
QY 1212 GGTGAGGTAGGCAAGGCTGCTGTGCTCTCAAGCTCTTGAGCAATGCTTACAT 1271
DB 874 GGAGATGTCGCAAGGCTGTGCTGCTGCTGCTGAACAAGCCGGTGCCTGCTGATTTGTG 933
QY 1272 ACCGAATGACCCCACTCTGCTCTGACAGGCTGCATGATGGTTCAGGTTGTTAAG 1331
DB 934 ACCGAGATTGACCTTATCTGTCTCTCCAGGCTACCATGGAAGGCTCCAGTCTTACT 993
QY 1332 CTAATGAGTCAATCCGCAAGTGCATGTCGTAATACTTGACAGGAATAGAAATGTA 1391
DB 994 CTAGAGGATGTCGTTCTGTGATGTTGATATCTTTGTCACACGACCGGTAACAGACAT 1053
QY 1392 GTGACACGGGAGCACTTGATCGCATGAAAAACAGTTGATGTCATATATGAGCCAC 1451
DB 1054 ATCATGTTGACCAATGAGGAAGATGAAGAACAATGCCATGTTTGGCAACATGCTCAC 1113
QY 1452 TCCAAACAGAAATGATGTGACACAGCCTCCGAC--TCCGAGCTGACGTGGAGCGA 1508
DB 1114 TTTGCAACGAATGCAATGCTTGTGCTGAGACCTACCTGCTGTCAGAGATTCACA 1173
QY 1509 GTAGCTTCTCAGGTGAGCATGTCATCTGCGCAGATGCAACAGAG--TTGTCTCTCTG 1565
DB 1174 ATTAAGCCTCAACACGACAGATGGGTCTTCCCTGACACCAACAGTGCATCATGCTTG 1233
QY 1566 GCAGAGGCTGCTCTACTCAATTTGAGCTGCTCCACAGTTCCACC--TTGTCTCTCTC 1622
DB 1234 GCTGAGGCTGCTCTACTCAATTTGAGCTGCTCCACAGACACCTAGTTTGTGATGTCG 1293
QY 1623 ATCAGAGCCACACAGAGCTTTGGCACTGATAGAACTTAATATGACC--CGAGGGG 1679

DB 1294 TGCTCGTCACTAACCAAGTCATTTGCCCACTGAGTGTGGAATGAAGAAGCAGTGGG 1353
QY 1680 CGATACAGCAGAGATGTACTGCTTCTCTAAGAAATGATGATACGTTGCCAGCTTG 1739
DB 1354 AAGTATGAGAAAGTGTATGCTTGTGCCAAACACCTGCAGCAGAGAGTGTGCTGACTT 1413
QY 1740 CATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTG 1799
DB 1414 CATCTCGAAAGCTCGAGCCAGCTTACCAAACTTTCGAAGATCAAGCTGACTACATT 1473
QY 1800 GGAATCAACAAAATGGGCCATTCAAACCTTAATTAACAGATAC 1844
DB 1474 AGCGTTCAGTTGAGGCTCTTACCAAGCCTGCTCACTACAGGTAC 1518

RESULT 6

US-08-930-894-1
Sequence 1, Application US/08930894
Patent No. 6037524
GENERAL INFORMATION:
APPLICANT: GREENLAND, Andrew James
APPLICANT: DRAPER, John
APPLICANT: SKIPSEY, Marc
APPLICANT: WARNER, Simon
TITLE OF INVENTION: S-ADENOSYL-L-HOMOCYSTEIN HYDROLASE PROMOTER
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pillsbury Madison & Sutro
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/930,894
FILING DATE: 09-OCT-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/00882
FILING DATE: 10-APR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9507381.3
FILING DATE: 10-APR-1995
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1767 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: SHH GENE FROM ASPARAGUS
FEATURE:
NAME/KEY: CDS
LOCATION: 26..1483
OTHER INFORMATION: /codon_start= 26
US-08-930-894-1

Query Match 11.6%; Score 214.6; DB 3; Length 1767;
Best Local Similarity 55.8%; Pred. No. 1.3e-46;
Matches 494; Conservative 0; Mismatches 379; Indels 12; Gaps 4;

QY 972 AAGTATCCAAAGCTGTTAAGAAGATCCGAGCATGTGGAAGAGAGCGTGAAGTGTGTT 1031
DB 596 AAGTACAGGAAGATGAAGATGATGATGTCGTTGTCGAGAGAGACACACCGGGGTC 655

QY	1032	CACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTGTTCCGGCCATGAAAGTCAT	1091
Db	656	AAGAGGCTTACCAGATGCAAGGCTAACAAATCCCTCTTTTCCCTGCGATCAATGTCAAT	715
QY	1092	GATTCGTATTACCAACAGAAAGTTTGATACTTGTAAGTCTGCCGAGAATCCATTTTGAT	1151
Db	716	GACTCCGTCACCAAGAGCAAGTTTGACAATCTGTATGATGCCGCACTCTCTTCCCGAT	775
QY	1152	GGCCTGAAGAGACACAGATGTGATGTTTGGTGGAAACAAGTGGTGGTGTGGCTAT	1211
Db	776	GGTCTGATGAGGGCCACTGATGTTATGATGTCTGGCAAGTTGCAGTTGTCTGCGGTTAT	835
QY	1212	GGTGAGGTAGGCAAGGCGCTGCTGCTGCTCTCAAAAGCTTTGAGCAATGTCTACAT	1271
Db	836	GGTGATGTCGGAGAGAGGCGTGTGCTGCTGCACCTCAAGCAGCGTGGTCCCGTATATGTG	895
QY	1272	ACCGAAATCGACCCCACTCTGTGCTCTGCAGGCGCTGCATGCAATGGGTTCAAGGTGTAAG	1331
Db	896	ACGGAGATCGACCCCACTCTGTGCTCTTCAAGCCCTAATGAGAGGCTTCAAGGTCTCACC	955
QY	1332	CTAAATGAAGTCATCCGGCAAGTCGATGTCGTAATTAAGTTCACAGGAATAAGAAATGTA	1391
Db	956	CTCGAGGATGTTGTCTCAGAGGCGGATATCTTGTATACCACCACCGGTAACAAAGGACATC	1015
QY	1392	GTGACACGGGAGCACTTGGATCGCATGAAACAGTTGTATCGTATGCAATATGGGCCAC	1451
Db	1016	ATCATGCTGGACCACATGAGAAAGATGAAAGAACAAATGCCATGTCTGCAACATTTGCTAC	1075
QY	1452	TCCAACACAGAAATCGAATGTGACCAGCCTCCGCAC--TCCGAGCTGACGTGGAGCGA	1508
Db	1076	TTTGACAACGAGATTGACATGCTAGGTTGGAGACATACCCTGGCATCAAGAGAATCACC	1135
QY	1509	GTACGTTCTCAGGTGGACCATGTTCATCTGGCCAGATGGCAA---ACGAGTGTCTCTCTG	1565
Db	1136	ATCAAGCCCCAGACTGACCCGGTGGGCTTCCCTGAAACCACACACTGGTATATTGTTCTT	1195
QY	1566	GCAGAGGGTGTCTACTCAATTTGAGCTGCTCCACAGTT---CCACCTTGTCTGTCTCC	1622
Db	1196	GCTGAGGGCCGACTCATGAACCTTGGGTGTGCCACTGTGTCACCCCAGCTTTGTCTATGTCC	1255
QY	1623	ATCACAGCCACACACAGCGTTTGGCACTGATAGAACTCTAATATG--CACCCGAGGGG	1679
Db	1256	TGCTCCTTCAACCAACGAGTGTATGCTCAGCTAGAGTTGTGGAATGAGAAAGCAAGCGGC	1315
QY	1680	CGATACAAGCAGATGTGTACTTGTCTTCTAAGAAATGATGAATACGTTGCCAGCTTG	1739
Db	1316	AAGTATGAGAAAGGTTTATCGTGTCTCCCAAGCATCTTGATGAGAAAGTAGCAGCGCTT	1375
QY	1740	CATCTGCCATCATTTGATGCCCACTTACAGAGCTGACAGATGACCAAGCAAAATATCTG	1799
Db	1376	CACCTTGGGCAAGCTCGGAGCCAAAGCTTACAAAGCTCAGCCCTTACACAGCGGACTACATC	1435
QY	1800	GGACTCAACAAAAATGGGCAATTCAACCTAATTATTACAGATAC	1844
Db	1436	AGCGTCCCCATCGAGGGTCCCTACAAAGCCACCTCACTACAGGTAC	1480

RESULT 7
US-08-896-005-2
Sequence 2, Application US/08896005
Patent No. 5854023
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Lal, Preeti
APPLICANT: Shah, Purvi
TITLE OF INVENTION: S¹ADENOSYL-L-HOMOCYSTEINE HYDROLASE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA

COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEO for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/896,005
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0337 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2226 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BLADTUT04
CLONE: 1519044

Query Match	6.38;	Score 117;	DB 2;	Length 2226;
Best Local Similarity	100.08;	Pred. No. 4.8e-21;		
Matches 117;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1728	GTGGCAGCTTGCATCTGCCATCATTTTGATGCCCCACCTTACAGAGCTGACAGATGACC	1787
	1	GTGGCAGCTTGCATCTGCCATCATTTTGATGCCCCACCTTACAGAGCTGACAGATGACC <td>60</td>	60
QY	1788	GCAAAATATCTGGGACTCACAAAAAATGGCCATTCAAACTAATTATTACAGATAC	
	61	GCAAAATATCTGGGACTCACAAAAAATGGCCATTCAAACTAATTATTACAGATAC	117
Db			

RESULT 8
 5-09-347-878-3
 Sequence 3, Application US/09347878C
 Patent No. 6376210
 GENERAL INFORMATION:
 APPLICANT: Yuan, Chong
 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES
 FILE REFERENCE: 25885-1651
 CURRENT APPLICATION NUMBER: US/09/347,878C
 CURRENT FILING DATE: 1999-07-06
 NUMBER OF SEQ ID NOS: 75
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 3

```

; LENGTH: 2226
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2226)
; OTHER INFORMATION: Polynucleotide encoding human
; OTHER INFORMATION: S-adenosyl-5-homocysteine hydrolase (SAHH) derived
; OTHER INFORMATION: from bladder; n-a, c, g, or t
; PUBLICATION INFORMATION:
; PATENT DOCUMENT NUMBER: 08/896,005
; PATENT FILING DATE: 1997-07-17
; PUBLICATION DATE: 1998-12-29
US-09-347-878-3

```

query match	6.38; score 117; db 4; length 2226;
-------------	-------------------------------------

Best Local Similarity 100.0%; Pred. No. 4.8e-21;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1728 GTGCCAGCTTGACATCGCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAA 1787
|||||
Db 1 GTGCCAGCTTGACATCGCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAA 60

QY 1788 GCAAAATATCTGGACATCAACAAAATGGCCCATCAACCTAATTATTACAGATAC 1844
|||||
Db 61 GCAAAATATCTGGACATCAACAAAATGGCCCATCAACCTAATTATTACAGATAC 117

RESULT 9

US-08-204-740-8

; Sequence 8, Application US/08204740
; Patent No. 5753432

GENERAL INFORMATION:

; APPLICANT: Gudkov, Andrei
; APPLICANT: Kazarov, Alexander
; APPLICANT: Mazo, Ilya
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods for Identifying Genetic
; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
; TITLE OF INVENTION: Growth in Cancer Cells
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allgretti & Witcoff, Ltd.
; STREET: 10 S. Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/204,740
FILING DATE: 04-MAR-1994
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: NO. 5753432nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 93,354-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:
LENGTH: 289 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA

US-08-204-740-8

Query Match 3.9%; Score 72.6; DB 1; Length 289;
Best Local Similarity 58.6%; Pred. No. 7.9e-10;
Matches 126; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 580 ACATCAAGCAGCAGAGATTGAGCGCCGGAGATTGAGATGACAGACAGACATGCTG 639
|||||
Db 75 ACATCGCCCTGGCTGCTGGGAGCGCAAGGCCCTGACATTTGCTGAGAGACGAGATGCCGG 134
QY 640 CTCTGATTTCACTCAGGAACGTGCTCAGGGGAGAGAGCCCTTGCTGCTAAATAG 699
|||||
Db 135 GCTGATGCGTATGCGGAGCGGTACTCGGCTTCCAGCCACTGAAGGGCCCCGCAATG 194
QY 700 TGGGCTGTACACACATCAGACCCAGACAGCGGTGTTGATTGAGACACTCTGTGCCCTGG 759
|||||
Db 195 CTGGCTGCGCTGCACATGACCGTGGAGAGCGGCCGCTCTCATTTGAGACCCCTGTCACCTGG 254

QY 760 GGGCTAGTCCCGCTGCTGCTGTGTAACATCTAC 794
|||||
Db 255 GTGCTGAGGTGACGTGCTGACGCTGCAACATCTTC 289

RESULT 10

US-09-081-167A-8

; Sequence 8, Application US/09081167A
; Patent No. 6083745

GENERAL INFORMATION:

; APPLICANT: Gudkov, Andrei
; APPLICANT: Kazarov, Alexander
; APPLICANT: Mazo, Ilya
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods for Identifying Genetic
; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
; TITLE OF INVENTION: Growth in Cancer Cells
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 S. Wacker Drive, 32nd Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/081,167A
FILING DATE: 18-MAY-1998
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: NO. 6083745nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 93,354-KK
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-0002

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:
LENGTH: 289 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA

US-09-081-167A-8

Query Match 3.9%; Score 72.6; DB 3; Length 289;
Best Local Similarity 58.6%; Pred. No. 7.9e-10;
Matches 126; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 580 ACATCAAGCAGCAGAGATTGAGCGCCGGAGATTGAGATGACAGACAGACATGCTG 639
|||||
Db 75 ACATCGCCCTGGCTGCTGGGAGCGCAAGGCCCTGACATTTGCTGAGAGACGAGATGCCGG 134
QY 640 CTCTGATTTCACTCAGGAACGTGCTCAGGGGAGAGAGCCCTTGCTGCTAAATAG 699
|||||
Db 135 GCTGATGCGTATGCGGAGCGGTACTCGGCTTCCAGCCACTGAAGGGCCCCGCAATG 194
QY 700 TGGGCTGTACACACATCAGACCCAGACAGCGGTGTTGATTGAGACACTCTGTGCCCTGG 759
|||||
Db 255 GTGCTGAGGTGACGTGCTGACGCTGCAACATCTTC 289

RESULT 11

US-09-081-395-8

Sequence 8, Application US/09081395

Patent No. 6083746

GENERAL INFORMATION:

APPLICANT: Gudkov, Andrei

APPLICANT: Kazarov, Alexander

APPLICANT: Mazo, Ilya

APPLICANT: Roninson, Igor B

TITLE OF INVENTION: Methods for Identifying Genetic

TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant

TITLE OF INVENTION: Growth in Cancer Cells

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff

STREET: 300 S. Wacker Drive, 32nd Floor

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/081,395

FILING DATE: 18-MAY-1998

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: No. 6083746nan, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 93,354-KK

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-913-0001

TELEFAX: 312-913-0002

TELEX:

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 289 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-09-081-395-8

Query Match

Best Local Similarity 3.9%; Score 72.6; DB 3; Length 289;

Matches 126; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 580 ACATCAAGCAGCAGAAATTTGGACGCCGGAGATTGAGATTGACAGACAGACATGCTG 639

DB 75 ACATCGGCGCTGCTGCTGGGAGCGCAAGGCCCTGACATGCTGAGAACGAGATGCCG 134

QY 640 CTCTGATTTCACTCAGGAACGTGCTCAGGGGAGAAAGCCCTTGCTGCTAAATAG 699

DB 135 GCCTGATGCGTATGCGGAGCGGTACTCGGCCCTCCAGCCACTGAAGGGCGCCGATCG 194

QY 700 TGGGCTGTACACATCAGACCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTG 759

DB 195 CTGGCTGCGCTGACATGACCGTGGAGACGGCCGCTCCTCATGAGACCCCTGTCACCCCTG 254

QY 760 GGGCTCAGTGCCGCTGCTGCTGCTGTAACATCTAC 794

DB 255 GTGCTGAGGTGACAGTGTCCAGCTGCAACATCTTC 289

RESULT 12

US-09-416-833-8

Sequence 8, Application US/09416833

Patent No. 6197521

GENERAL INFORMATION:

APPLICANT: Gudkov, Andrei

APPLICANT: Kazarov, Alexander

APPLICANT: Mazo, Ilya

APPLICANT: Roninson, Igor B

TITLE OF INVENTION: Methods for Identifying Genetic

TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant

TITLE OF INVENTION: Growth in Cancer Cells

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: Alligretti & Wilcoff, Ltd.

STREET: 10 S. Wacker Drive, Suite 3000

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/416,833

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/204,740

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: No. 6197521nan, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 93,354-C

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-715-1000

TELEFAX: 312-715-1234

TELEX: 910-221-5317

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 289 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-09-416-833-8

Query Match

Best Local Similarity 3.9%; Score 72.6; DB 4; Length 289;

Matches 126; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 580 ACATCAAGCAGCAGAAATTTGGACGCCGGAGATTGAGATTGACAGACAGACATGCTG 639

DB 75 ACATCGGCGCTGCTGCTGGGAGCGCAAGGCCCTGACATGCTGAGAACGAGATGCCG 134

QY 640 CTCTGATTTCACTCAGGAACGTGCTCAGGGGAGAAAGCCCTTGCTGCTAAATAG 699

DB 135 GCCTGATGCGTATGCGGAGCGGTACTCGGCCCTCCAGCCACTGAAGGGCGCCGATCG 194

QY 700 TGGGCTGTACACATCAGACCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTG 759

DB 195 CTGGCTGCGCTGACATGACCGTGGAGACGGCCGCTCCTCATGAGACCCCTGTCACCCCTG 254

QY 760 GGGCTCAGTGCCGCTGCTGCTGCTGTAACATCTAC 794

DB 255 GTGCTGAGGTGACAGTGTCCAGCTGCAACATCTTC 289

RESULT 13

PCT-US95-02521-8

Sequence 8, Application PC/TUS9502521

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: Methods for Identifying Genetic

TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant

TITLE OF INVENTION: Growth in Cancer Cells

NUMBER OF SEQUENCES: 13

```

: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/02521
: FILING DATE:
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 289 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: CDNA
: PCT-US95-02521-8

: Query Match 3.9%; Score 72.6; DB 5; Length 289;
: Best Local Similarity 58.6%; Pred. No. 7.9e-10;
: Matches 126; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 580 ACATCAAGCAGCAGAAATTGGACGCCGGAGATTGAGATTGCAGACGACATGTCG 639
   |||| | || | |||| | | |||| | | |||| | | |||| |
DB 75 ACATCGCGCTGGCTGCTCGGAGACGCAAGGCCCTGGACATTGCTGAGAACGAGATGCCG 134
   |||| | |||| | |||| | |||| | |||| | |||| | |||| |
QY 640 CTCGTATTTCACAGAAAGCTGCTCAGGGGGAGAGACCCCTGGCTGCTGCTAAATAG 699
   |||| | |||| | |||| | |||| | |||| | |||| | |||| |
DB 135 GCCTGATGCGTATGCGGAGGGGTACTCGGCTCCAGCCACTGAAGGGGCGCCGATCG 194
   |||| | |||| | |||| | |||| | |||| | |||| | |||| |
QY 700 TGGGCTGACACACATCACAGCCAGACAGCGGTGATTGAGACACTCTGCCCCG 759
   |||| | |||| | |||| | |||| | |||| | |||| | |||| |
DB 195 CTGGCTGCCCTGCACATGACCGTGAGAGACGGCCGCTCATTTGAGACCCCTGTCACCCCTGG 254
   |||| | |||| | |||| | |||| | |||| | |||| | |||| |
QY 760 GGGCTCAGTGGCGCTGCTGCTGCTTGTACATCTAC 794
   |||| | |||| | |||| | |||| | |||| | |||| | |||| |
DB 255 GTGCTGAGGTGCAAGTGTCTCAGCTGCAACATCTTC 289
   |||| | |||| | |||| | |||| | |||| | |||| | |||| |

RESULT 14
US-08-204-740-6
: Sequence 6, Application US/08204740
: Patent No. 5753432
: GENERAL INFORMATION:
: APPLICANT: Gudkov, Andrei
: APPLICANT: Kazarov, Alexander
: APPLICANT: Mazo, Ilya
: APPLICANT: Roninson, Igor B
: TITLE OF INVENTION: Methods for Identifying Genetic
: TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
: TITLE OF INVENTION: Growth In Cancer Cells
: NUMBER OF SEQUENCES: 13
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Allgrettl & Wiltcoff, Ltd.
: STREET: 10 S. Wacker Drive, Suite 3000
: CITY: Chicago
: STATE: Illinois
: COUNTRY: USA
: ZIP: 60606

: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/204,740
: FILING DATE: 04-MAR-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: No. 5753432nan, Kevin E
: REGISTRATION NUMBER: 35,303
: REFERENCE/DOCKET NUMBER: 93,354-C
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 312-715-1000

```

```

: TELEFAX: 312-715-1234
: TELEEX: 910-221-5317
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
:   LENGTH: 285 base pairs
:   TYPE: nucleic acid
:   STRANDEDNESS: single
:   TOPOLOGY: linear
:   MOLECULE TYPE: cDNA
: US-08-204-740-6

Query Match      3.7%; Score 68.4; DB 1; Length 285;
Best local similarity 56.8%; Pred. No. 9.9e-09;
Matches 126; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 573 GTGAGAACATCAAGCAGCAGAGATTGACGCCGAGAGATTGAGATTGCAGACCAAGAC 632
   || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 64 GTGCGGACATCGGACTGGCCGCTGGGAGCAGGAGCTCTGATATAGCTGAGATGAG 123
QY 633 ATGTCTGCTGTATTTCACATCAAGAAACGTGCTCAGGGGAGAGAGCCCTTGCTGTGCT 692
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 124 ATGCCAGGCTGATGCGCATGGCGGAGATGTACTCAAGCTCCAAAGCAGGAGGCTGCT 183
QY 693 AAAATAGTGGGCTGTACACACATCAAGCCAGACAGCGGTGTGATTGAGACACTCTGT 752
   || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 184 CGCATTCGCTGGCTGCTGCGCATGACCGGTGAGAGACTGCTGTCTCATTTGAGACTCTG 243
QY 753 GCCCTGGGGCTCAGTGCCTGCTGCTGTGTAACATCTAC 794
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 244 GCCCTGGGTGCTGAGGGCGGCTGCTCCAGCTGCAACATCTTC 285

RESULT 15
US-09-081-167A-6
: Sequence 6, Application US/09081167A
: Patent No. 6083745
: GENERAL INFORMATION:
:   APPLICANT: Gudkov, Andrei
:   APPLICANT: Kazarov, Alexander
:   APPLICANT: Mazo, Ilya
:   APPLICANT: Roninson, Igor B
:   TITLE OF INVENTION: Methods for Identifying Genetic
:   TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
:   TITLE OF INVENTION: Growth In Cancer Cells
:   NUMBER OF SEQUENCES: 15
:   CORRESPONDENCE ADDRESS:
:     ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
:     STREET: 300 S. Wacker Drive, 32nd Floor
:     CITY: Chicago
:     STATE: Illinois
:     COUNTRY: USA
:     ZIP: 60606
:   COMPUTER READABLE FORM:
:     MEDIUM TYPE: Floppy disk
:     COMPUTER: IBM PC compatible
:     OPERATING SYSTEM: PC-DOS/MS-DOS
:     SOFTWARE: PatentIn Release #1.0, Version #1.25
:     CURRENT APPLICATION DATA:
:       APPLICATION NUMBER: US/09/081,167A
:       FILING DATE: 18-MAY-1998
:       CLASSIFICATION: 435
:     ATTORNEY/AGENT INFORMATION:
:       NAME: No. 6083745nan, Kevin E
:       REGISTRATION NUMBER: 35,303
:       REFERENCE/DOCKET NUMBER: 93,354-KK
:     TELECOMMUNICATION INFORMATION:
:       TELEPHONE: 312-913-0001
:       TELEFAX: 312-913-0002
:     TELEX:
:   INFORMATION FOR SEQ ID NO: 6:
:     SEQUENCE CHARACTERISTICS:
:       LENGTH: 285 base pairs
:       TYPE: nucleic acid

```

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
us-09-081-167A-6

Query Match 3.7%; Score 68.4; DB 3; Length 285;
Best Local Similarity 56.8%; Pred. No. 9.9e-09;
Matches 126; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

```
QY 573 GTGAAGACATCAAGCAGGCAATTTGACCGCCGGAGATTGAGATTGCAGACAAGAC 632
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 64 GTCCGGGACATCGGACTGGCCCGCTGGGGAGCGGAAGGCTCTGATATAGCTGAGAATGAG 123

QY 633 ATGCTGCTCTGATTTCTACTCAGAAACGTGCTCAGGGGAGAGAGCCCTTGCTGCTGCT 692
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 124 ATGCCAGGTTGATGCGCATGCGGAGATGTACTCAGCCCTCCAAGCCACTGAAGGGTGCT 183

QY 693 AAAATAGTGGCTGTACACACATCACAGCCAGACAGCGGTGTGATTGAGACACTCTGT 752
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 184 CGCATGCTGGCTGCGCTGCGCATGACCGTGGAGACTGCTGTCTCATGTGAGACTCTCGTG 243

QY 753 GCCCTGGGGCTCAGTGCCGCTGGTCTGCTGTACATCTAC 794
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 244 GCCCTGGGTGCTGAGGCGGCTGCTCCAGCTGCAACATCTTC 285
```

Search completed: April 21, 2003, 05:16:14
Job time : 2569.43 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 02:39:45 ; Search time 369.111 Seconds
(without alignments)
5018.577 Million cell updates/sec

Title: US-09-782-051-1_COPY_1_1844

Perfect score: 1844
Sequence: 1 ggcgcggcagctcgagct.....aacctattattacagatc 1844

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 639749 seqs, 502280978 residues

Total number of hits satisfying chosen parameters: 1279498

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1844	100.0	2563	10 US-09-782-051-1	Sequence 1, Appl1
2	702.4	38.1	721	10 US-09-925-300-439	Sequence 439, App
3	461.4	25.0	553	10 US-09-867-701-2023	Sequence 2023, Ap
4	406.4	22.0	636	10 US-09-879-536-850	Sequence 850, App
5	402.4	21.8	2200	10 US-09-925-301-217	Sequence 217, App
6	402.4	21.8	2429	12 US-10-044-090-344	Sequence 344, App
7	337.5	18.3	384	10 US-09-880-107-346	Sequence 346, App
8	241.6	13.1	1461	10 US-09-759-990-1	Sequence 1, Appl1
9	225.6	12.2	1422	9 US-09-738-626-836	Sequence 836, App
10	225.6	12.2	1557	9 US-09-746-660A-97	Sequence 97, Appl
11	190	10.3	1877	12 US-10-021-121-1	Sequence 1, Appl1
12	186	10.1	1396	9 US-09-746-660A-101	Sequence 101, App
13	174	9.4	720	9 US-09-738-626-838	Sequence 838, App
14	172.4	9.3	3830	9 US-10-037-598-26	Sequence 26, Appl
15	172.4	9.3	513509	9 US-09-754-853A-4	Sequence 4, Appl1
16	168.2	9.1	433	10 US-09-960-352-4421	Sequence 4421, Ap
17	150.8	8.2	382	10 US-09-960-352-11499	Sequence 11499, A
18	138.4	7.5	284	10 US-09-920-300A-1423	Sequence 1423, Ap
19	138.4	7.5	284	12 US-10-033-528-1423	Sequence 1423, Ap

20	137.6	7.5	424	10 US-09-960-352-9160	Sequence 9160, Ap
21	136.8	7.4	288	9 US-09-736-457-1179	Sequence 1179, Ap
22	136.8	7.4	288	9 US-09-902-941-1179	Sequence 1179, Ap
23	136.8	7.4	288	9 US-09-849-626-1179	Sequence 1179, Ap
24	136.8	7.4	288	9 US-10-017-754-1179	Sequence 1179, Ap
25	136.4	7.4	288	10 US-09-998-598-2191	Sequence 2191, Ap
26	128.2	7.0	352	10 US-09-960-352-12928	Sequence 12928, A
27	124.4	6.7	283	10 US-09-960-352-7006	Sequence 7006, Ap
28	114.4	6.2	793	10 US-09-966-881-9	Sequence 9, Appl1
29	111.4	6.0	708	9 US-09-738-626-837	Sequence 837, App
30	108.2	5.9	389	10 US-09-960-352-10073	Sequence 10073, A
31	102.6	5.6	728	10 US-09-770-149-117	Sequence 117, App
32	100.6	5.5	237	10 US-09-960-352-14145	Sequence 14145, A
33	98.4	5.3	263	10 US-09-923-876-477	Sequence 477, App
34	96	5.2	410	10 US-09-960-352-1365	Sequence 1365, App
35	94.2	5.1	253	10 US-09-960-352-7284	Sequence 7284, App
36	91.4	5.0	399	10 US-09-878-574-838	Sequence 838, App
37	81.4	4.4	219	10 US-09-960-352-12043	Sequence 12043, A
38	79.2	4.3	456	10 US-09-960-352-2551	Sequence 2551, App
39	78.2	4.2	835	10 US-09-770-445-679	Sequence 679, App
40	75.2	4.1	405	10 US-09-960-352-5161	Sequence 5161, App
41	72.6	3.9	289	10 US-09-799-946-8	Sequence 8, Appl1
42	69.8	3.8	347	10 US-09-960-352-2035	Sequence 2035, App
43	68.4	3.7	285	10 US-09-799-946-6	Sequence 6, Appl1
44	67.6	3.7	163	10 US-09-878-574-9524	Sequence 9524, App
45	65.8	3.6	428	10 US-09-960-352-10555	Sequence 10555, A

ALIGNMENTS

RESULT 1
US-09-782-051-1
Sequence 1, Application US/09782051
Patent No. US20020035078A1
GENERAL INFORMATION:
APPLICANT: Hart, Derek N J
TITLE OF INVENTION: Enzyme having S-adenosyl-L-homocysteine hydrolase
TITLE OF INVENTION: (AHCY) type activity
FILE REFERENCE: 24305 MRB
CURRENT APPLICATION NUMBER: US/09/782,051
CURRENT FILING DATE: 2001-02-14
PRIOR APPLICATION NUMBER: PCT/NZ97/00133
PRIOR FILING DATE: 1997-10-06
PRIOR APPLICATION NUMBER: NZ 299507
PRIOR FILING DATE: 1996-10-04
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 2563
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (3)..(1847)
OTHER INFORMATION: Open reading frame extends without a stop codon
OTHER INFORMATION: for the full 5' nucleotide sequence. The
OTHER INFORMATION: Initiation codon has yet to be identified.
US-09-782-051-1

Query Match	100.0%;	Score 1844;	DB 10;	Length 2563;
Best local Similarity	100.0%;	Pred. No. 0;		
Matches 1844;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GGCGGGGCGAGCTCGAGCTGCTCTCTGTTCTCTTGCGCCACCGTGGCT	60	
Db	1	GGCGGGGCGAGCTCGAGCTGCTCTCTCTGTTCTCTTGCGCCACCGTGGCT	60	
QY	61	GTCGGGCTGCTTGGGCTGCGACAGACAGCGGTGGCCACAGCCTCAGAGCCGA	120	
Db	61	GTCGGGCTGCTTGGGCTGCGACAGACAGCGGTGGCCACAGCCTCAGAGCCGA	120	
QY	121	CGCAGCTGACGAGGCGCGGCGAGGAGGCGATGCGGTGCGAGGCGCGCGCGC	180	

```
Db 121 CGCAGCTCGACGAGGGGGGGGAGAGGGTGGCGGATCGCGTGTGAGAGGGGGGGGG 180
QY 181 GGGCAGGGGGGGGGGGGGGGGAGAGGGGGGAGAGGGGGGGGGGGGGGGGGGGGGGGGG 240
Db 181 GGGCAGGGGGGGGGGGGGGGGAGAGGGGGGAGAGGGGGGGGGGGGGGGGGGGGGGGGG 240
QY 241 GGGGGGGGGGGGAGATGTCGATGCTGACGGGATGCCGCTGCGGGGGGGGGGGGGGGGG 300
Db 241 GGGGGGGGGGGGAGATGTCGATGCTGACGGGATGCCGCTGCGGGGGGGGGGGGGGGGG 300
QY 301 TGAAGCAGGGCCAGAGATCGAGAGCGCCGAGAGTACTCCTTCATGCGCCACCGTCACCA 360
Db 301 TGAAGCAGGGCCAGAGATCGAGAGCGCCGAGAGTACTCCTTCATGCGCCACCGTCACCA 360
QY 361 AGGGCCCAAGAGCAAAATCCAGTTGCTGATGACATGACAGAGTTCCACAAATTCGCCA 420
Db 361 AGGGCCCAAGAGCAAAATCCAGTTGCTGATGACATGACAGAGTTCCACAAATTCGCCA 420
QY 421 CCAGAACTGGCCGAGATCTTGTCTGCTGATCTCAGAGTCTCCTCCACTGACAGCTACA 480
Db 421 CCAGAACTGGCCGAGATCTTGTCTGCTGATCTCAGAGTCTCCTCCACTGACAGCTACA 480
QY 481 GTTCAGCTGCATCTACACAGATAGCTGTGATGATGAGTTTCTCCCGAGAGAGCAGC 540
Db 481 GTTCAGCTGCATCTACACAGATAGCTGTGATGATGAGTTTCTCCCGAGAGAGCAGC 540
QY 541 AAACCACTCCAGAGGGGAGAGCAATTTCTGTGAAGAATCAAGCAGGAGAAATTTG 600
Db 541 AAACCACTCCAGAGGGGAGAGCAATTTCTGTGAAGAATCAAGCAGGAGAAATTTG 600
QY 601 GACGCGGGAGATGAGATTGACAGAGCAAGATGCTGCTGATTTCACTCAGGAAC 660
Db 601 GACGCGGGAGATGAGATTGACAGAGCAAGATGCTGCTGATTTCACTCAGGAAC 660
QY 661 GTGCTCAGGGGAGAGAGCCCTTGCTGCTGCTAAATAGTGGGCTGTACACATCACAG 720
Db 661 GTGCTCAGGGGAGAGAGCCCTTGCTGCTGCTAAATAGTGGGCTGTACACATCACAG 720
QY 721 CCCAGACAGCGGTGTGATTGAGACACTGCTGCTGCGGGCTCAGTGGCTGTG 780
Db 721 CCCAGACAGCGGTGTGATTGAGACACTGCTGCTGCGGGCTCAGTGGCTGTG 780
QY 781 CTGTACATCTACTCAACTCAGATGAGATGAGTGCAGCACTGGCTGAGGCTGAGTTG 840
Db 781 CTGTACATCTACTCAACTCAGATGAGATGAGTGCAGCACTGGCTGAGGCTGAGTTG 840
QY 841 CAGTGTGCTGTGAGAGGGGAGTCAAGATGACTTCTGTTGTTGATGACCGCTGTG 900
Db 841 CAGTGTGCTGTGAGAGGGGAGTCAAGATGACTTCTGTTGTTGATGACCGCTGTG 900
QY 901 TGAACATGATGGGTGGCAGGCCAATGATCTGATGATGGGGGAGACTTAACCCACT 960
Db 901 TGAACATGATGGGTGGCAGGCCAATGATCTGATGATGGGGGAGACTTAACCCACT 960
QY 961 GGGTTTATAGAGTATCCAAACGTTTAAAGAGATCCGAGGCAATGGAAGAGAGCG 1020
Db 961 GGGTTTATAGAGTATCCAAACGTTTAAAGAGATCCGAGGCAATGGAAGAGAGCG 1020
QY 1021 TGAAGTGTGTCACAGGCTGTATCAGCTTCCAAAGCTGGGAGAGCTGTGTCGGCCA 1080
Db 1021 TGAAGTGTGTCACAGGCTGTATCAGCTTCCAAAGCTGGGAGAGCTGTGTCGGCCA 1080
QY 1081 TGAACGTCAATGATCTGTTACCAACAGAGTTGATTAATGTTGCTGCTGCGAGAA 1140
Db 1081 TGAACGTCAATGATCTGTTACCAACAGAGTTGATTAATGTTGCTGCTGCGAGAA 1140
QY 1141 CCATTTTGAATGGCTGAAGAGACACAGATGATGTTGGTGGAAACAAGTGTGG 1200
Db 1141 CCATTTTGAATGGCTGAAGAGACACAGATGATGTTGGTGGAAACAAGTGTGG 1200
QY 1201 TGTGTGCTATGCTGAGTAGGCAAGGGCTGTGCTGCTCAAAAGCTTTGAGACAA 1260
Db 1201 TGTGTGCTATGCTGAGTAGGCAAGGGCTGTGCTGCTGCTCAAAAGCTTTGAGACAA 1260
```

```
Db 1201 TGTGTGCTATGCTGAGTAGGCAAGGGCTGTGCTGCTGCTCAAAAGCTTTGAGACAA 1260
QY 1261 TTGTCTACATTAACCAATCGAACCCCATCTGCTCTGACAGGCTGATGATGGTTCA 1320
Db 1261 TTGTCTACATTAACCAATCGAACCCCATCTGCTCTGACAGGCTGATGATGGTTCA 1320
QY 1321 GGGTGTAAAGCTAAATGAGTATCCGGCACTGATGTCGTAATTAATTCACAGAA 1380
Db 1321 GGGTGTAAAGCTAAATGAGTATCCGGCACTGATGTCGTAATTAATTCACAGAA 1380
QY 1381 ATAGAAATGTAGTACACAGGAGCACTTGATGTCATGAAAAACAGTTGATGTCATGA 1440
Db 1381 ATAGAAATGTAGTACACAGGAGCACTTGATGTCATGAAAAACAGTTGATGTCATGA 1440
QY 1441 ATATGGCCACCTCAACACAGAAATGATGTGACAGCCCTCCGACCTCGAGCTGACGT 1500
Db 1441 ATATGGCCACCTCAACACAGAAATGATGTGACAGCCCTCCGACCTCGAGCTGACGT 1500
QY 1501 GGGAGCAGTACGTTCTCAGGTGAGTGCATGCTGCGCAGATGGCAAGAGAGTTGTC 1560
Db 1501 GGGAGCAGTACGTTCTCAGGTGAGTGCATGCTGCGCAGATGGCAAGAGAGTTGTC 1560
QY 1561 TCCTGGCAGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTCCACCTTTGTTCTGT 1620
Db 1561 TCCTGGCAGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTCCACCTTTGTTCTGT 1620
QY 1621 CCATCACAGCCCAACACAGAGCTTTGGCAGTGTAGAACTCTATAATGACCCGAGGGG 1680
Db 1621 CCATCACAGCCCAACACAGAGCTTTGGCAGTGTAGAACTCTATAATGACCCGAGGGG 1680
QY 1681 GATACAGCAGAGTGTACTTCTGCTCTTAAGAAATGATGATAGTACGTTGCCAGCTGC 1740
Db 1681 GATACAGCAGAGTGTACTTCTGCTCTTAAGAAATGATGATAGTACGTTGCCAGCTGC 1740
QY 1741 ATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1800
Db 1741 ATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1800
QY 1801 GACTCAACAAAATGGGCCATTCAACCTTAATTAACAGATAC 1844
Db 1801 GACTCAACAAAATGGGCCATTCAACCTTAATTAACAGATAC 1844
```

RESULT 2
US-09-925-300-439

; Sequence 439, Application US/09925300
; Patent No. US20020151681A1

; GENERAL INFORMATION:

; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

; FILE REFERENCE: PA101

; CURRENT APPLICATION NUMBER: US/09/925,300

; CURRENT FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: PCT/US00/05988

; PRIOR FILING DATE: 2000-03-08

; PRIOR APPLICATION NUMBER: 60/124,270

; NUMBER OF SEQ ID NOS: 1890

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 439

; LENGTH: 721

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (688)

; OTHER INFORMATION: n equals a,t,g, or c

; US-09-925-300-439

Query Match 38.1%; Score 702.4; DB 10; Length 721;
Best Local Similarity 98.0%; Pred. No. 2.9e-209;
Matches 703; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

```
OY 1 GGGCGGGGAGGTCGGAGCTCGAGCTGCTGCTTCTGTTCTCTTGTGGCCACCCTGCT 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 4 GGGCGGGGAGGTCGGAGCTCGAGCTGCTGCTTCTGTTCTCTTGTGGCCACCCTGCT 63
OY 61 GTCCGGCTGCTTGGGCTGCCGAACAGAGGGCTGGGCCACAGCAGCTCAGAAAGCCGA 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 64 GTCCGGCTGCTTGGGCTGCCGAACAGAGGGCTGGGCCACAGCAGCTCAGAAAGCCGA 123
OY 121 CGCAGCTCGAGCGAGGGGCGGCGGAGAGGGTGGCGATCGGTGTCGAGGGGCGCGCGC 180
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 124 CGCAGCTCGAGCGAGGGGCGGCGGAGAGGGTGGCGATCGGTGTCGAGGGGCGCGCGC 183
OY 181 GGGCAGGGGCGGCGGCGGAGAGGGGAGAGGGGCGGCGGCGGCGGCGGCGGCGGCGG 240
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 184 GGGCAGGGGCGGCGGCGGAGAGGGGAGAGGGGCGGCGGCGGCGGCGGCGGCGGCGG 243
OY 241 GGGCGGGGCGGGAATGTCGATGCTGACGCGATGCCGCTGCCCGGGGTCGGGAGAGAGC 300
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 244 GGGCGGGGCGGGAATGTCGATGCTGACGCGATGCCGCTGCCCGGGGTCGGGAGAGAGC 303
OY 301 TGAAGCAGGGCCAGAGAGATCGAGGACGCCGAGAACTACTCTTCATGGCCACCCTCACC 360
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 304 TGAAGCAGGGCCAGAGAGATCGAGGACGCCGAGAACTACTCTTCATGGCCACCCTCACC 363
OY 361 AGGCGCCCAAGAGCAATCCAGTTGCTGATGACATGACAGAGTTCACCAATTCGCCCA 420
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 364 AGGCGCCCAAGAGCAATCCAGTTGCTGATGACATGACAGAGTTCACCAATTCGCCCA 423
OY 421 CCAAACTGGCCGAAGATCTTTGTCGCTGATCTCAGATCTCAGTCCCTCAGTACAGCTACA 480
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 424 CCAAACTGGCCGAAGATCTTTGTCGCTGATCTCAGATCTCAGTCCCTCAGTACAGCTACA 483
OY 481 GTTCAGCTGCATCTACACAGATAGCTGATGATGAGGTTTCTCCCGAGAGAAAGCAGC 540
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 484 GTTCAGCTGCATCTACACAGATAGCTGATGATGAGGTTTCTCCCGAGAGAAAGCAGC 543
OY 541 AAACCACTCCAAAGGGCAGCAGCAATTTCTGTGAAGAACATCAAGCAGAGAGAAATTG 600
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 544 AAACCACTCCAAAGGGCAGCAGCAATTTCTGTGAAGAACATCAAGCAGAGAGAAATTG 603
OY 601 GACGCGGGGAGATTGAGATTGACAGCAGACATGTCTGTGATTTCACTCAGGAAAC 660
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 604 GACGCGGGGAGATTGAGATTGACAGCAGACATGTCTGTGATTTCACTCAGGAAAC 663
OY 661 GTGCTCAGGGGAGAGCCCTTGCTGCTAAATAGTGCGGTGACACACATCA 717
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 664 GTGCTCAGGGGAGAGCCCTTGCTGCTAAATAGTGCGGTGACACACATCA 720
```

RESULT 3

```
US-09-867-701-2023
; Sequence 2023, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aqlate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2023
; LENGTH: 553
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(553)
; OTHER INFORMATION: n - A,T,C or G
```

US-09-867-701-2023

```
Query Match 25.0%; Score 461.4; DB 10; Length 553;
Best Local Similarity 97.38; Pred. No. 6.2e-134;
Matches 510; Conservative 0; Mismatches 9; Indels 5; Gaps 4;
OY 1322 GGTGTAAGCTAAATGAAGTCATCCGGCAAGTCGATGCTGTAATACCTTGACAGGAAA 1381
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 GGTGTAAGCTAAATGAAGTCATCCGGCAAGTCGATGCTGTAATACCTTGACAGGAAA 60
OY 1382 TAAGATGTAGTGACACGGGAGCAGCAGTGGATGCGATGAAAAACAGTTGTATGTAACAA 1441
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 TAAGATGTAGTGACACGGGAGCAGCAGTGGATGCGATGAAAAACAGTTGTATGTAACAA 120
OY 1442 TATGGCCACTCCAACAGAAATGATGTGACAGCCTCCGACTCCGAGCTGACGTG 1501
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 TATGGCCACTCCAACAGAAATGATGTGACAGCCTCCGACTCCGAGCTGACGTG 180
OY 1502 GGAGCAGTACGTTCTCAGGTGACCATGTCTGACCAGATGGCAACGAGTGTCT 1561
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 181 GGAGCAGTACGTTCTCAGGTGACCATGTCTGACCAGAT-GNAAACGAGTGTCTCT 239
OY 1562 CCTGGCAGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTCCACACTTGTCTGTC 1621
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 240 CCTGGCAGAGGGTCTCTACTCAATTTGAGCTGCTCCACAGTCCACACTTGTCTGTC 299
OY 1622 CATCACAGCCACAACAGAGCTTTGGCAGCTGATGAACTCTATAATGACCCGAGGGCG 1681
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 300 CATCACAGCCACAACAGAGCTTTGGCAGCTGATGAACTCTATAATGACCCGAGGGCG 359
OY 1682 ATACAAGCAGGATGTGTAATCTGCTCTCTTAAGAAA-TGGATGAATACGTTGCGAGCTGC 1740
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 360 ATACAAGCAGGATGTGTAATCTGCTCTCTTAAGAAA-TGGATGAATACGTTGCGAGCTGC 419
OY 1741 ATCTGCCATCATTTGATGCCCACTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 1800
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 420 ATCTGCCATCATTTGATGCCCACTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG 478
OY 1801 GACTCAACAAAAATGGCCCATTCAAACCTAATTAATACAGATAC 1844
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 479 GACTCAACAAAA-TGGGCCATTCAACCTAATTAATTAACAGATAC 520
```

RESULT 4

```
US-09-879-536-850/c
; Sequence 850, Application US/09879536
; Patent No. US20020144298A1
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dertl, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/879,536
; CURRENT FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/088,801
; PRIOR FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 850
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Homo sapiens
```

FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(636)
OTHER INFORMATION: n - A,T,C or G
US-09-879-536-850

Query Match 22.0%; Score 406.4; DB 10; Length 636;
Best Local Similarity 99.5%; Pred. No. 1.1e-116;
Matches 418; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 707 TACACATACACAGCCAGACAGC-GGTGTGATTGAGACACTCTGTGCCCCGGGGCTC 765
DB 420 TACACATACACAGCCAGACAGCAGCGGTGTGATTGAGACACTCTGTGCCCCGGGGCTC 361
QY 766 AGTCCGCTGCTGCTGTGTACATCTACTCACTCAGATGAGTAGCTGACACCTGG 825
DB 360 AGTCCGCTGCTGCTGTGTACATCTACTCACTCAGATGAGTAGCTGACACCTGG 301
QY 826 CTGAGGCTGAGTGTGAGTGTGCTTGAAGGCGAGTCAGAGATGACTTCTGTGTGT 885
DB 300 CTGAGGCTGAGTGTGAGTGTGCTTGAAGGCGAGTCAGAGATGACTTCTGTGTGT 241
QY 886 GTATTGACCGCTGTGTGAACATGATGGGTGGAGCCCAACATGATCTGTGATGGG 945
DB 240 GTATTGACCGCTGTGTGAACATGATGGGTGGAGCCCAACATGATCTGTGATGGG 181
QY 946 GAGACTTAACCCACTGGGTTTATAGAGTATCCAAACGTGTTTAAAGAGATCCGAGCA 1005
DB 180 GAGACTTAACCCACTGGGTTTATAGAGTATCCAAACGTGTTTAAAGAGATCCGAGCA 121
QY 1006 TTGTGGAAGAGAGCGTGTGCTGTACAGGCTGTATCAGCTCTCCAAAGTGGGAAGC 1065
DB 120 TTGTGGAAGAGAGCGTGTGCTGTACAGGCTGTATCAGCTCTCCAAAGTGGGAAGC 61
QY 1066 TCTGTGTCGGGCATGAACGTCAATGATCTGTACCAACAGAGTTGATTAATTGT 1125
DB 60 TCTGTGTCGGGCATGAACGTCAATGATCTGTACCAACAGAGTTGATTAATTGT 1

RESULT 5

US-09-925-301-217
Sequence 217, Application US/09925301
Patent No. US20020052308A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA106
CURRENT APPLICATION NUMBER: US/09/925,301
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05882
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 217
LENGTH: 2200
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2188)
OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-217

Query Match 21.8%; Score 402.4; DB 10; Length 2200;
Best Local Similarity 58.4%; Pred. No. 3.8e-115;
Matches 741; Conservative 0; Mismatches 521; Indels 6; Gaps 2;

QY 580 ACATCAAGCAGCAGATTTGGACGCGCGGAGATGAGATTGACAGAGCAAGACATGCTG 639
DB 108 ACATCGGCGCTGCTGCTGGGAGCGAAGCGCTGACATTTGCTGAGACGAGATGCCGG 167

QY 640 CTCTGATTTCACTCAGAAACGTGCTCAGGGGAGAGACCCCTTGCTGTGCTAAATAG 699
DB 168 GCCGTGATGCGTATGCGGAGCGGTACTCGGCCCTCAAGCCACTGAAGGGCGCCGATCG 227
QY 700 TGGGCTGTACACATCAACAGCCAGACAGCGGTGTGATTGAGACACTCTGTGCCCTGG 759
DB 228 CTGGCTGCTGCACATGACCGGTGAGAGCGGCCCTCTCTATTGAGACCCCTGCTACCCCTGG 287
QY 760 GGGCTCAGTGGCCGCTGCTGCTGCTGTAACATCTACTCACTCAGATGAGTAGTGTGAG 819
DB 288 GTGCTGAGGTGAGTGTGCTGCTGCTGCAACATCTCTCCACCAGGACCATGCGGGGCTG 347
QY 820 CACTGCTGAGGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 879
DB 348 CCATTGCCAAGGCTGCGATTCGCTGTATGCTGTGAGGCGGAAACGAGAGAGTACC 407
QY 880 GGTGCTGATTGACCGCTGTGTGAACATGATGGGTGGAGCCCAACATGATCTGTGATG 939
DB 408 TGTGCTGATGAGCAGACCCCTGTACTTCAAGAGCGGGCCCTCAACATGATCTGTGAGC 467
QY 940 ATGGGGAGACTTAACCCACTGGGTTTATAGAGTATCCAAACGTGTTTAAAGATCC 999
DB 468 ACGGGGCGACCTCACCACCTCATCCACACCAAGTACCCGACGCTTCTGCCAGCATCC 527
QY 1000 GAGGCATGTGGAAGAGACCGTGTGCTGTACAGGCTGTATCAGCTCTCCAAAGCTG 1059
DB 528 GAGGCATCTGTGAGAGACACCGACTGGGTGCTCACAACCTCTTACAAGATGATGGCAATG 587
QY 1060 GGAAGCTCTGTGTTCCGGCCATGAACGTCAATGATCTGTGTTACCAACAGAGTTGATA 1119
DB 588 GGATCCTCAAGGTGCTGCTGCTCAATGATCTGTGTTACCAACAGAGTTGATA 647
QY 1120 ACTGTACTGCTGCGGAGATTCATTTGATGCGCTGAAGAGACACAGATGATGT 1179
DB 648 ACCTCTATGCTGCGGAGATTCCTCATAGATGGATCAAGCGGCACAGATGATGA 707
QY 1180 TTGTGGAACAAGT 1239
DB 708 TTGTGGAACAAGT 767
QY 1240 CTCTCAAGCTCTTGAGCAATTTGTCTACATTTACCAATCGAACCCCTGTGCTCTGC 1299
DB 768 CCCTGCGGGGTTTGGAGCCCGGCTCATCATCAGAGATTGACCCCATCAAGCAGCTGC 827
QY 1300 AGGCTGATGATGGGTTTCAAGGTGTGTAAGCTTAATGAAGTCAATCCGGCAGTGTG 1359
DB 828 AGGCTGATGATGGGTTTCAAGGTGTGTAAGCTTAATGAAGTCAATCCGGCAGTGTG 887
QY 1360 TCGTATTAAGTGTGCAAGAAATAGATGTAGTACACAGGAGCAGTGTGATGCAATGA 1419
DB 888 TCTTTGACACACAGAGCTGTATTGACATCACTTGGCCGACCTTGGACAGATGA 947
QY 1420 AAAACAGTTGTATGCAATATATGAGGCGACCTCAACACAGAAATGATGACCAAGC 1479
DB 948 AGGATGATGCAATGTGTATGATGAGTGTGACACTTGAAGTGTGATGATGATGATG 1007
QY 1480 TCCGCACTCCGAGCTGACGTGGAGCGAGTGTCTCAGGTGAGACCATGTCTGTG 1539
DB 1008 TCAACGAGAACCGCTGGAAGAGTGAACATCAACCCGAGGTGAGCCGCTGTGATG 1067
QY 1540 CAGATGGCAAGAGTGTCTCTCTGCGAGAGAGGTGTCTACTCAATTTGAGCTGTCCA 1599
DB 1068 AGATGGGCGCGCATCATCTGCTGCGGAGGCTGTGCTCAACCTGTGCTGTGCA 1127
QY 1600 CAG--TTCCACCTTGTCTGTCTCATCATCAGCCCAACACAGGCTTGGCACTGATG 1656
DB 1128 TGGCCACCCAGCTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1187
QY 1657 AACTCTAATATGACCCGAGGGGCGATACAGAGATGTACTTGTCTCTAAGAA 1716
DB 1188 AGCTGTGACCAATCCAGACAAG--TACCCGCTGGGGTTCATTTCTGCCAAGAAGC 1244
QY 1717 TTGATGATATGTTGCCAGCTTGCATCTGCCATCATTTGATGCCCACTTACAGAGCTGA 1776

Db 1245 TGGATGAGGAGGCTGGAGAGCCACCTGGGCAAGCTGATGTGAAGTGTGACCAAGCTAA 1304
QY 1777 CAGATGACCAAGCAAAATATCTGGGACTCAACAAAAATGGCCATTCAAACCTAATTAT 1836
Db 1305 CTGAGAGCAAGCCCACTACCTGGGATGTCTGTGATGGCCCTTCAGAGCCGGATCACT 1364
QY 1837 ACAGATAC 1844
Db 1365 ACCGCTAC 1372

RESULT 6

US-10-044-090-344
; Sequence 344, Application US/10044090
; Patent No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044, 090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 344
; LENGTH: 2429
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 1468237CB1
US-10-044-090-344

Query Match 21.8%; Score 402.4; DB 12; Length 2429;
Best Local Similarity 58.4%; Pred. No. 4e-115;
Matches 741; Conservative 0; Mismatches 521; Indels 6; Gaps 2;

QY 580 ACATCAAGCAGGCAAGATTGGACGCCGGAGATTGAGATGACAGCAAGACATGCTG 639
Db 108 ACATCGGCTGCTGCTGGGAGCGCAAGGCCCTGGACATGCTGAGAACGAGATGCCG 167
QY 640 CTCTGATTTCACTCAGAAAGCTGCTCAGGGGAGAGCCCTTGCTGCTAAATAG 699
Db 168 GCCTGATGCGTATGCGGAGCGGTACTCGGCTCCAAGCCACTGAAGGGCGCCGATCG 227
QY 700 TGGGCTGTACACATCACAGCCCAAGACGCGTGTGATGAGACACTCTGTGCCCTG 759
Db 228 CTGGCTGCTGACATGACCGGTGAGAGCGCGCTCTCATGAGACCCCTGTCACCCCTG 287
QY 760 GGGCTCAGTCCGCTGCTGCTGCTGTAACATCTACTCACTCAGAAATGAAGTGTGAG 819
Db 288 GTGCTGAGGTGAGTGTGCTCAAGCTGCAACATCTTCTCCAGGAGCAACCATGCGGCTG 347
QY 820 CACTGGCTGAGGCTGAGTGTGCTGCTGCTGGAAGGGGAGTCAAGATGACTTCT 879
Db 348 CCATGCGCAAGGCTGCAATGCGGTGTATGCTGGAAGGGGCAAGCGAGAGATACC 407
QY 880 GGTGCTGTATGACCGCTGTGTAACATGATGGTGGCAGGCCAATGATCTGATG 939
Db 408 TGTGCTGATGAGCAGACCCCTGATCTCAAGAGCGGCCCTCAACATGATCTGAGC 467
QY 940 ATGGGGAGACTTAACCCAGGCTTTATAAGAGTATCCAAACGTGTTAAGAGATCC 999
Db 468 ACGGGGGGCACTCACCACCTTCATCCACCAAGTACCCGAGCTTCTGCCAGGATCC 527
QY 1000 GAGGCTGTGAGAGAGAGCGGTGCTGCTCAGAGGCTGTATCAGCTCTCCAAAGCTG 1059
Db 528 GAGGCTGTGAGAGAGAGCGGTGCTGCTCAGAGGCTGTATCAGCTCTCCAAAGCTG 587
QY 1060 GGAAGCTCTGTGCTGGGCTGAGAGAGCTCAATGATCTGTTACCAAGAGATTGATA 1119
Db 588 GGAAGCTCTGAGGCTGCTGAGAGAGAGCTCAATGATCTGTTACCAAGAGATTGATA 647

QY 1120 ACTTGTACTGCTGCCGAGATTCATTTTGATGGCCCTGAAGAGGACACAGATGTGAT 1179
Db 648 ACCCTATGCTGCCGAGAGCTCCATATGATGGCATCAAGCGGCCACAGATGTGATGA 707
QY 1180 TTGTTGGGAAACAAGTGTGTGTGTGCTATGTTGAGTGTAGGCAAGGCTGTGCTG 1239
Db 708 TTGCGGCAAGGTAGCGGTGTGAGAGGCTATGTTGATGTGGGCAAGGCTGTGCCAG 767
QY 1240 CTCTCAAGCTCTTGAGCAATGCTACATTTACCGAATCGACCCCATCTGTCTGTC 1299
Db 768 CCTGCGGGGTTTCGAGAGCCCGCTCATCATCACCAGATGTGACCCCATCAAGCATTC 827
QY 1300 AGGCTGCTGATGATGGTTCAGGCTGTAAAGCTAAATGAGTCAATCCGCAAGTGTG 1359
Db 828 AGGCTGCTGATGAGGCTGTATGATGATGATGATGATGATGATGATGATGATGATG 887
QY 1360 TCGTAATTAATTCACAGGAAATAGATGTGATGATGATGATGATGATGATGATGAT 1419
Db 888 TCTTGTACACACACAGGCTGTATGATGATGATGATGATGATGATGATGATGATGAT 947
QY 1420 AAAACAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1479
Db 948 AGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1007
QY 1480 TCCGCACTCCGAGCTGAGCTGAGGAGGAGTACGTTCTCAGTGTGAGCATGTCATG 1539
Db 1008 TCAACGAGAACGCGCTGAGAGAGTGAACATCAAGCCGAGTGTGATGATGATGATG 1067
QY 1540 CAGATGCAAAAGAGTGTCTCTCTGCGAGAGGCTGCTCTCTCTCTCTCTCTCTCT 1599
Db 1068 AGAATGGGCGCGCATCATCTCTGCTGCGAGGCTGCTGCTGCTGCTGCTGCTGCT 1127
QY 1600 CAG--TTCCACCTTTGTTCTGTCATCAGACGACCAACACAGGCTTTGACATGAT 1656
Db 1128 TGGGCAACCCAGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATG 1187
QY 1657 AACTCTAATGACACCCGAGGCGATCAAGCAGATGTGATGATGATGATGATGATGAT 1716
Db 1188 AGCTGTGAGCCATCCAGACAAG--TACCCGTTGGGTTCAATTTCTGCCCCAAGAAG 1244
QY 1717 TGGATGAATACGTTGCCAGCTTGCATGCTGATTTGATGCGCACCTTACAGAGCTGA 1776
Db 1245 TGGATGAGGAGTGGCTGAAGCCCACTGGGCAAGCTGATGTGATGATGATGATGAT 1304
QY 1777 CAGATGACCAAGCAAAATATCTGGGACTCAACAAAAATGGCCATTCAAACCTAATT 1836
Db 1305 CTGAGAGCAAGCCCACTACCTGGGATGTCTGTGATGCGCCCTTCAAGCCGATGACT 1364
QY 1837 ACAGATAC 1844
Db 1365 ACCGCTAC 1372

RESULT 7

US-09-880-107-346
; Sequence 346, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 346

LENGTH: 384
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA157401
US-09-880-107-346

Query Match 18.3%; Score 337; DB 10; Length 384;
Best Local Similarity 99.0%; Pred. No. 4.2e-95;
Matches 381; Conservative 0; Mismatches 0; Indels 4; Gaps 4;

QY 992 GAAGATCCGAGGCGATGTGGAAGAGAGAGCGGTGACTGGTGTTCACAGGCTGTATCAGCTCTC 1051
DB 1 GAAGATCCGAGGCGATGTGGAAGAGAGAGCGGTGACTGGTGTTCACAGGCTGTATCAGCTCTC 60
QY 1052 CAAAGCTGGGAAGCTCTGTGTCCGGCCATGAACGTCATGATTTCTGTTACCAACAGAA 1111
DB 61 CAAAGCTGGGA-CTGTGTGTCCGGCCATGAACGTCATGATTTCTGTTACCAACAGAA 119
QY 1112 GTTGATACTGTGCTGCTCCGAGAAATCCATTTGATGCGCTGAAGAGACCAAGAGA 1171
DB 120 GTTGATACTGTGCTGCTCCGAGAAATCCATTTGATGCGCTGAAGAGACCAAGAGA 179
QY 1172 TGTGATGTTGGTGGGAACAAGTGTGTGTGCTATGCTGAGTAGGCAAGGCTG 1231
DB 180 TGTGATGTTGGTGGGAACAAGTGTGTGTGCTATGCTGAGTAGGCAAGGCTG 239
QY 1232 -CTGTGCTGCTCTC-AAAGCTCTTGAGCAATGTCTACATTACCGAATCGACCCATC 1289
DB 240 CCTGTGCTGCTCTCAAAAGCTCTTGAGCAATGTCTACATTACCGAATCGACCCATC 299
QY 1290 TGTGCTCTGAGGCGCTGCATGATGGTTCAGGCTGT-AAAGCTAAATGAAGTCAATCCG 1348
DB 300 TGTGCTCTGAGGCGCTGCATGATGGTTCAGGCTGTAAAGCTAAATGAAGTCAATCCG 359
QY 1349 GCAAGTCGATGTCTTAATACTTGC 1373
DB 360 GCAAGTCGATGTCTTAATACTTGC 384

RESULT 8

US-09-759-990-1
Sequence 1, Application US/09759990
Patent No. US20020119491A1
GENERAL INFORMATION:
APPLICANT: Anticancer, Inc.
APPLICANT: Xu, Mingxu
TITLE OF INVENTION: HIGH EXPRESSION AND PRODUCTION OF HIGN
TITLE OF INVENTION: SPECIFIC ACTIVITY RECOMBINANT S-ADENOSYLHOMOCYSTAINASE
TITLE OF INVENTION: (SAH) AND IMPROVED ASSAYS FOR S-ADENOSYLMETHIONINE (SAM)
FILE REFERENCE: 31276-20026.00
CURRENT APPLICATION NUMBER: US/09/759,990
PRIOR FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: US 60/176,444
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1461
TYPE: DNA
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Nucleotide sequence encoding SAH
US-09-759-990-1

Query Match 13.1%; Score 241.6; DB 10; Length 1461;
Best Local Similarity 57.8%; Pred. No. 6.2e-65;
Matches 490; Conservative 0; Mismatches 349; Indels 9; Gaps 3;

QY 1002 GGCATGTGGAAGAGAGCGTGAAGTGTGTTCACAGGCTGTATCAGCTCTCAAGCTGGG 1061
DB 613 GGTGTTCCGAAGAGACAACAACAGGTGTCCACCGCTTACCAAGCTCGAAGAGAGGGC 672

QY 1062 AAGCTCTGTCTCCGGCCATGAAGCTCAATGATTTCTGTTACCAACAGAGTTGATTAAC 1121
DB 673 AACTCTCTCTTCCAGCCATCAACGTCAACGACGCTGTTACAAAGTCCAAAGTTCGATTAAC 732
QY 1122 TTGTAAGCTGCTCCGAGAAATCCATTTGATGGCTGAAGAGAGACCAAGATGTGATGTT 1181
DB 733 ATCTAAGGCTGCGCCACTCCCTTATCGATGATGATCAACCGTGTCCGATGATGATC 792
QY 1182 GGTGGAAACAAGTGTGTGTGTGCTATGCTGAGTAGGCAAGGCTGCTGTGCTGCT 1241
DB 793 GGGGCAAGACAGCTCTGTCATGAGGTTACGGGATGTCCGCAAGGCTGCGCTCAATCC 852
QY 1242 CTCAAAGCTCTTGAGCAATGTCTACATTACCAATCAACCCATCTGTGCTCTGAG 1301
DB 853 CTCGCTGGCCAAAGGCTGCGCTTATCATCACAGAACTCGACCCCAATCTGCGCTCTCAG 912
QY 1302 GCGTCATGATGGGTTACAGGTTGTTAAAGCTAAATGAAGTCAATCCGCAAGTCAATC 1361
DB 913 GCTGCCATGGAAGCTTACAGGTTCCGCGCATCGAGGAAGTGTCAAGGATGTGATATC 972
QY 1362 GTAATACTTGACAGGAATAAAGATGATGACACGGGACCACTTGATCGCATGAA 1421
DB 973 TTGCTTACATGCACAGGAATGCGATATCATCTGTGACATGATGCGCCAGATGAG 1032
QY 1422 AACAGTTGATGCTATGATGATGAGGCTCCAAACAGAAATGATGTGACCAAGCTC 1481
DB 1033 GATAAGGCTATTGTGCTAATGCGCCACTTGATTAAGAAATGATATACAGATGGCTC 1092
QY 1482 CGCA--CTCCGAGAGCTGAGTGGAGCGAGTACGTTCTCAGGTGACCATGTCACTGG 1538
DB 1093 ATGAATACCCAGGATCAACACATCCCAATCAAGCCAGAAATACGACATGTGGAAATC 1152
QY 1539 CCAGATGGCAACGAGTGTCTCTCTGCGACAGAGGCTGTCTACTCAATTTGAGCTGCTC 1598
DB 1153 CCAGATGGCCACGCTATCTCTCTCTGCTGAGGCGCTTCAACCTTGGCTGCGCT 1212
QY 1599 ACAGT---CCACCTTGTCTGTCTCATCACAGCCACAACAGGCTTTGGCACTGATA 1655
DB 1213 ACAGTCAACCATCTTCTGTATGTCAATGTCTATCAACCAAGACACTGCTCAGCTC 1272
QY 1656 GAAGCTTATATGACCCGAGGCGGATACAGCAAGATGTGTACTTGTCTTCTAAGAA 1715
DB 1273 GACCTCTACGA--AAAGAGAGGAATCTCGAGAGAAAGTTTACACACTTCCGAAGCAT 1329
QY 1716 ATGATGAATACGTGCGACCTTGCACTTGCACTATTTGATGCCACCTTACAGAGCTG 1775
DB 1330 CTCGATGAAGAGTCTGCTGCGCTCCACCTCGGATCTCTGATGTCCACCTTCAAAAGCT 1389
QY 1776 ACAGATGACCAAGAAATATCTGGAGCTCAACAAATGGGCAATCAACCTAATAT 1835
DB 1390 ACACAGAGCAGGCTGACTACATCAAGCTTCAAGTTGAGGCTCTTACAGTCTGATGCT 1449
QY 1836 TACAGATA 1843
DB 1450 TACCGTTA 1457

RESULT 9

US-09-738-626-836
Sequence 836, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:

APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO

```

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 836
; LENGTH: 1422
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-836

```

Query Match	12.28;	Score 225.6;	DB 9;	Length 1422;
Best Local Similarity	56.88;	Pred. No. 6.2e-60;		
Matches 477; Conservative	0;	Mismatches 354;	Indels 9;	Gaps 3;

OY	1011	GAAGAGAGCGTGACTGTGTTGCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTTGT	1070
Db	586	GAGGAACCACCAACCGGTGTGCACCGCGCTGTACCACTTCGCTGAAGAAGGCGTGCCT	645
OY	1071	GTTCCGGCCATGAACGTCAATGATTCTGTATTACCAACAAGAAGTTGATACTTGTACTGC	1130
Db	646	TTCACAGCATGAACGTCAACGACGCTGTACACCAAGTCCAAGTTTGATAACAAGTAACGCC	705
OY	1131	TGCCGAGATCCATTTTGATGGCCTGAAGAGACACAGATGTGATGTTGGTGGAAA	1190
Db	706	ACCCGCCACTCCCTGATCGACGGCATCAACCGCGCACATGACATGCTCATGGCGGCAAG	765
OY	1191	CAGTGGTGTGTGTGCTATGGTGAGGTAGGCAAGGCGCTGCTGTGCTCTCAAGCT	1250
Db	766	AACGTGCTGTCTGGGTTACGGCGATGTCCGCAAGGCGCTGCGCTGAGGCTTTCAGACGGC	825
OY	1251	CTTGAGCAATGTCTACATTAACCGAATCGACCCATCTGTGCTCTGCAGGCGCTGCATG	1310
Db	826	CAGGCGCTCGCTCAAGGTCAACCGAGCTGACCCCAATCAACGCTCTTCAGGCTGTATG	885
OY	1311	GATGGGTTCAAGGTGTAAAGCTAAATGAAGTCATCCGCAAGTGCATGTCTGAATAACT	1370
Db	886	GATGGCTACTGTGTGTCAACGTTGATGAGGCCATCGAAGACGCCGACATCGTGAATCAC	945
OY	1371	TGCACAGGAATAAGATGTAGTGACACAGGAGACACTTGATCCGATGAAAAACAAGTTGT	1430
Db	946	GCGACCCGGCAACAAGACATCATTTCTTCGAGCAAGTAGCTCAAGATGAAGATCACGCT	1005
OY	1431	ATCGTATGCAATATGGGCGCACTCCAACACAGAAATCGATGTG--ACCAAGCTCCGCACT	1487
Db	1006	CTGCTGGGCAACATCGGTCACTTTGATAATGAGATCGATATGCATTCCCTGTTCACCGC	1065
OY	1488	CCGGAGCTGACGTGGGAGCGAGTACGTTCTCAGGTGACCATGTCATCTGGCCAGATGGC	1547
Db	1066	GACGACGTCAACCGCACAGATCAAGCCACAGGTGACAGATTCAACCTTCTCCACCGGT	1125
OY	1548	AAACGAGTTGTCTCTCGGAGAGGGTCTACTCTCAATTGAGCTGCTCCAC--AGTT	1604
Db	1126	CGCTCCATCATCGTCTCTCCGAAGGTGCGCTGTTGAACCTTGCAACGCCACCGGACAC	1185
OY	1605	CCCACCTTTGTTCTGTCAATCACAGCCACACACAGGCTTTGGCACTGATAGAATCTAT	1664
Db	1186	CCATCATTTGTATGTCCAACTCTTTCGCCGATCAGACCAATTGGCAGATCGAACCTGTC	1245
OY	1665	AATGCACCCGAGGGCGATACAAAGCAGATGTGTACTTGTCTCTTAAGAAATGATGA	1724
Db	1246	CA---AAACGAAGACAGTACGAGAAGCAGAGGTCTACCGCTGTGCTTAAGGTTCTCGACGA	13022
OY	1725	TACGTTGCCAGCTTGCATCTGCCATCAATTGATGCCACCTTACAGAGCTGACAGATGAC	1784
Db	1303	AAGGTGGCAGCATCCACGTTGAGGCTCTCGGCGGTACAGTCAACCGAATGACCAAGGAG	13622

QY 1785 CAAGCAAAATATCTGGGACTCAACAAAAATGGGCCATTCAAAACCTAATTATTACAGATTAC 1844

Db 1363 CAGGCTGAGTACATCGGCGTTGACGTTGCAGGCCCATTCGAAGCCGGAGCAGTACCGGCTAC 1422

RESULT 10
US-09-746

```

; Sequence 97, Application US/09746660A
; Publication No. US20030049804A1
; GENERAL INFORMATION:

```

```

: APPLICANT: Pompejus, Markus
: APPLICANT: Kroger, Burkhard
: APPLICANT: Schroder, Hartwig
: APPLICANT: Zelder, Oskar
: APPLICANT: Habershauer, Gregor
: APPLICANT: Kim, Jun-Won
: APPLICANT: Lee, Heung-Schlick
: APPLICANT: Hwang, Byung-Joon
: TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
: TITLE OF INVENTION: METABOLIC PATHWAY PROTEINS.
: FILE REFERENCE: BGI-121CP2

```

```

: CURRENT APPLICATION NUMBER: US/09/746,660A
: CURRENT FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 09/606740
: PRIOR FILING DATE: 2000-06-23
: PRIOR APPLICATION NUMBER: 09/603124
: PRIOR FILING DATE: 2000-06-23
: PRIOR APPLICATION NUMBER: 60/141031
: PRIOR FILING DATE: 1999-06-25
: PRIOR APPLICATION NUMBER: 60/142101
: PRIOR FILING DATE: 1999-07-02
: PRIOR APPLICATION NUMBER: 60/148613
: PRIOR FILING DATE: 1999-08-12
: PRIOR APPLICATION NUMBER: 60/187970
: PRIOR FILING DATE: 2000-03-09
: PRIOR APPLICATION NUMBER: DE 19931420.9
: PRIOR FILING DATE: 1999-07-08
: NUMBER OF SEQ ID NOS: 125
: SOFTWARE: PatentIn Vers. 2.0
: SEQ ID NO 97
: LENGTH: 1557

```

```

; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(1534)
; OTHER INFORMATION: RXN00132
;
DS-09-746--660A-97

```

Query Match	12.2%;	Score 225.6;	DB 9;	Length 1557;
Best Local Similarity	56.8%;	Pred. No. 6.5e-60;		
Matches 477; Conservative	0;	Mismatches 354;	Indels 9;	Gaps 3

OY	1011	GAAGAGAGCGTGA	CTGTGTTTCACAGGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGT	1070
Db	698	GAGNAACCACCACCGGTGTGCACCGCTGTACCACTTCGCTGAAGAAGGCGTGCTGCT	757	
OY	1071	GTTCCGCCCATGAA	CGTCAATGATTCTGTTACCAAACAGAGTTTGATAACTTGTA	1130
Dd	758	TTCACAGCGATGA	ACGTCAACGAGCTGTCCACCAAGTCCAGTTTGATAACAAGTACGGC	817
OY	1131	TGCCGAGATCCAT	TTTGGATGGCCGTGAAGAGGACCACAGATGTGATGTTTGGTGGAAA	1190
Dd	818	AACCGCCACTCCC	TGATCGACGGCATCAACCGGCCACTGACATGCTCATGGGGCGCAAG	877
OY	1191	CAAGTGGTGGTGT	TGGCTATGTTGAGGTAGGCAGAAGGCGTGTGCTGCTCTCAAGCT	1250
Dd	878	AACGTGCTGTCT	CGGTTACGGCGCATGTCCGCCAAGGCGCTGAGGCTTTCGACGGC	937
OY	1251	CTTGAGCAATTGT	CTACATTACCGAAATCGACCCCATCTGTGCTCTGCAGGCGCTGCATG	1310
Dd	938	CAGGGCGCTCGCT	CAAGGTACACCGAAGCTGACCCCAATCAACGCTCTTCAGGCTCTGATG	997

OY	1311	GATGGGTTCAAGGTGCTGAAGAAGCTAAATGAGTCAATCCGGCCAAGTCGATGTCTGAATAACT	1370
Dd	998	GATGGCTACTCTGTGTGTACCCGTTGATGAGGCCATCGAGGACGCCGACATCGTGATCACC	1057
OY	1371	TGCACAGGMAATAGAAATGTAGTAGCACACGGGAGCACCTTGGATCGCATGAAAAACAGTTGT	1430
Dd	1058	GCGAACCGGCACAAGGACATCATTTCTTCGAGCAGATGCTCAAGATGAAGATCACGCT	1117
OY	1431	ATCGTATGCATATATGGGCCACTCCAACACAGAAATCGATGTG--ACCAGCCTCCGCACT	1487
Dd	1118	CTGCTGGGCACATCGGTCACTTTGATTAATGAGATCGATATGCAFTTCCCCTGTTGCACGC	1177
OY	1488	CCGGAGCTGACGCTGGGAGCGAGTACGTTCTCAGGTGACCATGTCATCTGGCCAGATGGC	1547
Dd	1178	GACGACGTCAACCCGACACGATCAAGCCACAGGTGACAGAGTTCAACCTTCTCCACCGGT	1237
OY	1548	AAACGAGTTGTCTCCTCTGGCAGAGGGTCTCTACTCAATTGAGCTGCTCCAC--AGTT	1604
Dd	1238	CGCTCCATCATCTGCTCTGTCCGAAGTGCCTGTGTAACCTTGCAACGCCACCGGACAC	1297
OY	1605	CCCCACCTTGTCTGTCCATCACAGCCACAACACAGGCTTTGGCACTGATAGAACTCTAT	1664
Dd	1298	CCATCATTTGTCTATGTCCAACTCTTTCGCCGATCAAGCAATTTGGCAGATCGAATCTTC	1357
OY	1665	AATGCACCCGAGGGCGCATACAAGCAGGATGTGTACTTGCTTCCCTAAGAAATGATGAA	1724
Dd	1358	CA---AAACGAAGAGACGTACGAGAAGAGGTCTACCGTCTGCCTAAGGTTCTCGACGA	1414
OY	1725	TACGTTGCCAGCTTGCATCTGCCATCATTTGTATGCCCACTTACAGAGCTGACAGATGAC	1784
Dd	1415	AAGTGGCACGCACTCCAGCTTGAGGCTCTCGGGGTTCAGCTCACCGAACTGACCAAGAG	1474
OY	1785	CAGCAAAATATCTGGGACTCAACAAAATGGGCCATTCAAACCTAATTATTACAGATAC	1844
Dd	1475	CAGGCTGAGTACATCGGCGTTGACGTTGCAGGGCCCAATTCAGCCGAGACACTACCGTAC	1534

RESULT 11
US-10-021-121-1/c
; Sequence 1, Application US/10021121
; Patent No. US2002014244A1

APPLICANT: Caras, Ingrid W
 TITLE OF INVENTION: A2-1 Neurotrophic Factor
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genentech, Inc.
 STREET: 1 DNA Way
 CITY: South San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94080
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 Inch, 1.44 MB floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Winpatin (Genentech)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/021,121
 FILING DATE: 06-Dec-2001
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/635,130
 FILING DATE: 19-Mar-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Torchia, PhD., Timothy E.
 REGISTRATION NUMBER: 36,700
 REFERENCE/DOCKET NUMBER: P1001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650/225-8674
 TELEFAX: 650/952-9881
 INFORMATION FOR SEQ ID NO: 1:

```

; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1877 base pairs
;   TYPE: Nucleic Acid
;   STRANDEDNESS: Double
;   TOPOLOGY: Linear
;
; FEATURE:
;   NAME/KEY: Extra Cellular Domain
;   LOCATION: 244-899
;   IDENTIFICATION METHOD:
;   OTHER INFORMATION:
;
; FEATURE:
;   NAME/KEY: Transmembrane Domain
;   LOCATION: 901-978
;   IDENTIFICATION METHOD:
;   OTHER INFORMATION:
;
; FEATURE:
;   NAME/KEY: signal peptide
;   LOCATION: 244-321
;   IDENTIFICATION METHOD:
;   OTHER INFORMATION:
;
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-021-121-1

```

Query Match	10.3%	Score 190;	DB 12;	Length 1877;
Best Local Similarity	100.0%	Pred. No. 9.8e-49;		
Matches 190; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1283	CCCCATCTGTGCTCTCGACAGCCCTGCATGGATGGGGTTCAGGGTGTTAAAGCTAATAAGAAT	1342
Dd	1446	CCCCATCTGTGCTCTCGACAGGCCCTGCATGGATGGGGTTCAGGGTGTTAAAGCTAATAAGAAT	1387
QY	1343	CATCCGGCAAGTCGATGTCGTATAACTTGCACAGGAATAAGATGTAGTAGCACCGGA	1402
Dd	1386	CATCCGGCAAGTCGATGTCGTATAACTTGCACAGGAATAAGATGTAGTAGCACCGGA	1327
QY	1403	GCACTTGGATCGCATGAAAAACAAGTTGTATCGTATGCAATATGGGCCACTCCAACACAGA	1462
Dd	1326	GCACTTGGATCGCATGAAAAACAAGTTGTATCGTATGCAATATGGGCCACTCCAACACAGA	1267
QY	1463	AATCGATGTG 1472 	
Dd	1266	AATCGATGTG 1257	

RESULT. 12.
US-09-746-660A-101
; Sequence 101, Application US/09746660A
; Publication No. US20030049804A1

APPLICANT: Pompejus, Markus
 APPLICANT: Kroger, Burkhard
 APPLICANT: Schroder, Hartwig
 APPLICANT: Zelder, Oskar
 APPLICANT: Haberhauer, Gregor
 APPLICANT: Kim, Jun-Won
 APPLICANT: Lee, Heung-Schick
 APPLICANT: Hwang, Byung-Joon
 TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
 TITLE OF INVENTION: METABOLIC PATHWAY PROTEINS
 FILE REFERENCE: BGI-12ICP2
 CURRENT APPLICATION NUMBER: US/09/746,660A
 CURRENT FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 09/606740
 PRIOR FILING DATE: 2000-06-23
 PRIOR APPLICATION NUMBER: 09/603124
 PRIOR FILING DATE: 2000-06-23
 PRIOR APPLICATION NUMBER: 60/141031
 PRIOR FILING DATE: 1999-06-25
 PRIOR APPLICATION NUMBER: 60/142101
 PRIOR FILING DATE: 1999-07-02
 PRIOR APPLICATION NUMBER: 60/148613
 PRIOR FILING DATE: 1999-08-12
 PRIOR APPLICATION NUMBER: 60/187970

;; PRIOR FILING DATE: 2000-03-09.
;; PRIOR APPLICATION NUMBER: DE 19931420.9
;; PRIOR FILING DATE: 1999-07-08
;; NUMBER OF SEQ ID NOS: 125
;; SOFTWARE: Patentln Vers. 2.0
;; SEQ ID NO 101
;; LENGTH: 1396
;; TYPE: DNA
;; ORGANISM: Corynebacterium glutamicum
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (101)..(1396)
;; OTHER INFORMATION: FRXA01371
US-09-746-660A-101

Query Match 10.1%; Score 186; DB 9; Length 1396;
Best Local Similarity 57.3%; Pred. No. 1.5e-47;
Matches 377; Conservative 0; Mismatches 275; Indels 6; Gaps 2;

QY 1011 GAAGAGAGCGTGAAGTGTGTTACAGAGCGTGTATCAGCTCTCCAAAGCTGGAGACTCTGT 1070
DB 698 GAGGAACACCAACACCGGTGTGACACCGCTGTACCACTCGCTGAAGAGCGCTGCT 757
QY 1071 GTCCGGCCATGAACGTCATGATCTGTATACCAACAAGAGTTGATTAATCTTACTGC 1130
DB 758 TTCCAGCGATGAACGTCACAGCGCTGTACCAAGTCCAAAGTTGATTAACAGTACGGC 817
QY 1131 TGCCGAGATCCATTTTGATGCGCTGAAGAGGACACAGATGTGATGTTGGGAAA 1190
DB 818 ACCCGCACCTCCCTGATGACGCGCATCAACCGCCACTGACATGCTCATGGCGGCAAG 877
QY 1191 CAAGTGTGTGTGTGCTATGTGAGGTAGGCAAGGGCTGTGCTGTCTCAAAAGCT 1250
DB 878 AACGTGTGTGTGTGCGGTACGCGATGTGCGCAAGGGCTGTGAGGCTTTCGACGGC 937
QY 1251 CTGAGCAATGTCTACATTACGAAATGCAACCCCATCTGTGCTGACGGCTGCATG 1310
DB 938 CAGGGCGCTGCGTCAAGTCAACCGAAGCTGACCAATCAACGCTCTTCAAGCTGTATG 997
QY 1311 GATGGGTCAAGGGTGTAAAGCTAAATGATCATCCGGCAAGTGCATGCTAATAACT 1370
DB 998 GATGGCTACTCTGTGTGTCACCGTGTGATGAGCCATCGAAGACCGGACATCGTATCAC 1057
QY 1371 TGCACAGAAATAAGATGTAGTACACGGGACACTTGGATCGCATGAAACAAGTTGT 1430
DB 1058 GCGACCGGCAACAAGACATCATTTCTTCGAGCAGATGCTCAAGATGAAGATCACGCT 1117
QY 1431 ATCGTATGCAATATGGGCCACTCCAACACAGAAATCGATGTGACCAAGCTC---CGCACT 1487
DB 1118 CTGCTGGGCAACATCGGTCACTTTGATATGAGATCGATATGCAATCCCTGTGACCGC 1177
QY 1488 CCGGAGCTGACGTGGGAGCGAGTACGTTCTCAGGTGAGCCATGTCTGCGCAGATGGC 1547
DB 1178 GACGACGTACCCGACACGATCAAGCCAGGTGAGAGTTCACTTCTCCACCGGT 1237
QY 1548 AAACGAGTGTCTCTCTGCGAGAGGGTCTACTCAATTGAGCTGCTCCACAGTTCC 1607
DB 1238 CGCTCATCATCTGCTCTGCGAAGGTGCTGTGAACCTTGCAACGCCACCGGACAC 1297
QY 1608 ACC---TTGTCTCTGTCATCACAGCCACACACAGGCTTGGCACTGATGAAGTCT 1662
DB 1298 CCATTTGTCATGTCCAACTCTTTCCGCGATCAGACCATTTGGCGAGATCGAACTGT 1355

RESULT 13
US-09-738-626-838/c
; Sequence 838, Application US/09738626
; Patent No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO

;; APPLICANT: OCHIAI, KEIKO
;; APPLICANT: YOKOI, HARUHIKO
;; APPLICANT: TATEISHI, NAOKO
;; APPLICANT: SENOH, AKIHIRO
;; APPLICANT: IKEDA, MASATO
;; APPLICANT: OZAKI, AKIO
;; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;; FILE REFERENCE: 249-125
;; CURRENT APPLICATION NUMBER: US/09/738,626
;; PRIOR FILING DATE: 2000-12-18
;; PRIOR APPLICATION NUMBER: JP 99/377484
;; PRIOR FILING DATE: 1999-12-16
;; PRIOR APPLICATION NUMBER: JP 00/159162
;; PRIOR FILING DATE: 2000-04-07
;; PRIOR APPLICATION NUMBER: JP 00/280988
;; PRIOR FILING DATE: 2000-08-03
;; NUMBER OF SEQ ID NOS: 7059
;; SOFTWARE: Patentln ver. 3.0
;; SEQ ID NO 838
;; LENGTH: 720
;; TYPE: DNA
;; ORGANISM: Corynebacterium glutamicum
US-09-738-626-838

Query Match 9.4%; Score 174; DB 9; Length 720;
Best Local Similarity 61.0%; Pred. No. 5.9e-44;
Matches 282; Conservative 0; Mismatches 180; Indels 0; Gaps 0;

QY 1011 GAAGAGAGCGTGAAGTGTGTTACAGAGCGTGTATCAGCTCTCCAAAGCTGGAGACTCTGT 1070
DB 465 GAGGAACACCAACACCGGTGTGACACCGCTGTACCACTCGCTGAAGAGCGCTGCT 406
QY 1071 GTCCGGCCATGAACGTCATGATCTGTATACCAACAAGAGTTGATTAATCTTACTGC 1130
DB 405 TTCCAGCGATGAACGTCACAGCGCTGTACCAAGTCCAAAGTTGATTAACAGTACGGC 346
QY 1131 TGCCGAGATCCATTTTGATGCGCTGAAGAGGACACAGATGTGATGTTGGGAAA 1190
DB 405 TTCCAGCGATGAACGTCACAGCGCTGTACCAAGTCCAAAGTTGATTAACAGTACGGC 346
QY 1191 CAAGTGTGTGTGTGCTATGTGAGGTAGGCAAGGGCTGTGCTGTCTCAAAAGCT 1250
DB 345 ACCCGCACCTCCCTGATGACGCGCATCAACCGCCACTGACATGCTCATGGCGGCAAG 286
QY 1251 CTTGAGCAATGTCTACATTACGAAATGCAACCCCATCTGTGCTGACGGCTGCATG 1310
DB 285 AACGTGTGTGTGCGGTACGCGATGTGCGCAAGGGCTGTGAGGCTTTCGACGGC 226
QY 1311 GATGGGTCAAGGGTGTAAAGCTAAATGATCATCCGGCAAGTGCATGCTAATAACT 1370
DB 225 CAGGGCGCTCGGTCAGGTCAACCGAAGCTGACCAATCAACGCTCTTCAAGCTCTGATG 166
QY 1371 GATGGGTCAAGGGTGTAAAGCTAAATGATCATCCGGCAAGTGCATGCTAATAACT 1370
DB 165 GATGGCTACTCTGTGTGTCACCGTGTGATGAGGCCATCGAAGACCGGACATCGTATCAC 106
QY 1371 TGCACAGAAATAAGATGTAGTACACGGGAGCACTTGGATCGCATGAAACAAGTTGT 1430
DB 105 GCGACCGGCAACAAGACATCATTTCTTCGAGCAGATGCTCAAGATGAAGATCACGCT 46
QY 1431 ATCGTATGCAATATGGGCCACTCCAACACAGAAATCGATGTG 1472
DB 45 CTGCTGGCAACATCGGTCACTTTGATATGAGATCGATATG 4

RESULT 14
US-10-037-598-26
; Sequence 26, Application US/10037598
; Patent No. US20020157143A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Co
; APPLICANT: Monsanto Co
; APPLICANT: Delaney, Xavier
; TITLE OF INVENTION: Soybean Plants with Enhanced Yields and Methods for Breeding f
; TITLE OF INVENTION: Screening of Soybean Plants with Enhanced Yields
; FILE REFERENCE: 38-21(52175)B
; CURRENT APPLICATION NUMBER: US/10/037,598

Tue Apr 22 10:05:26 2003

us-09-782-051-1_copy_1_1844.rnpb

Page 11

OY 1739 GCATCTGCCCATCATTTTGATTGCCCCACCCCTTACAGAGCTGCAGATGACCCAAGCAAATAACT 1798
| | | | | | | | | | | | | | | |
DB 178318 TCACCTGGGCAAACTTGGAGCCTAAGCTGACCCAGCAGCTTAGCAAGTCGCCAGCGGTGATTACAT 178259

OY 1799 GGGACTCAACAAAAATGGGCCATTCAAACCTAATTAATTACAGATAC 1844
| | | | | | | | | | | | | | | |
DB 178258 CAGTGTGCCCTGTTGAGGGTCCATACAAGCCCTGCTCACTACAGGTAC 178213

Search completed: April 21, 2003, 07:45:05
Job time : 1204.11 secs